

COHERENT OPTICAL ARRAY RECEIVER FOR PPM SIGNALS UNDER ATMOSPHERIC TURBULENCE

Thesis by
Michela Muñoz Fernández

In Partial Fulfillment of the Requirements for
the Degree of

Doctor of Philosophy



CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2006

(Defended December 19, 2005)

© 2006

Michela Muñoz Fernández

All Rights Reserved

To my mother Alicia, for all her immense love,
encouragement, kindness, and support.

Thesis Committee

Babak Hassibi

Associate Professor of Electrical Engineering

Victor Vilnrotter

Digital Signal Processing Group, Jet Propulsion Laboratory

Ryan Mukai

Digital Signal Processing Group, Jet Propulsion Laboratory

Charles Elachi

Vice President; Director of Jet Propulsion Laboratory; Professor of Electrical Engineering and Planetary Science

Hideo Mabuchi

Associate Professor of Physics and Control and Dynamical Systems

Demetris Psaltis

Thomas G. Myers Professor of Electrical Engineering

Changhuei Yang

Assistant Professor of Electrical Engineering and Bioengineering

Acknowledgements

First of all, I wish to thank my advisor at Caltech, Prof. Babak Hassibi for welcoming me as part of his group and for his sincere mentorship and advice, and my advisors at the Jet Propulsion Laboratory, Dr. Victor A. Vilnrotter and Dr. Ryan Mukai, for their guidance and support in making it possible to conduct this research.

I am very grateful to the members of my defense and candidacy committees for their advice and helpful suggestions: Prof. Charles Elachi, Prof. Hideo Mabuchi, Prof. Demetris Psaltis, and Prof. Changhuei Yang.

I extend my appreciation to the Jet Propulsion Laboratory (JPL) IND (Interplanetary Network Directorate) Advanced Program and staff for their dedication to the higher education of research scientists and engineers. The financial support and honor of working with Dr. Tsun-Yee Yan, and Dr. Norm Lay's Group is gratefully acknowledged. The research described in this dissertation was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA).

Special thanks to the members of the Digital Signal Processing Group at JPL, especially to Dr. Marvin Simon for his invaluable advice that has enhanced this research. I thank immensely Carlos Esproles for teaching me the experimental skills required for my research; I really enjoyed working all the countless hours in the optical laboratory. I am very grateful to Dr. Hammid Hemmati for allowing me to use the laboratory facilities

of his group and to Bill Farr for his technical advice, and for letting me use equipment that has been very useful for my research.

I also acknowledge Dr. Marc Rayman for giving me the opportunity to accomplish my dream to come to JPL and work on an incredibly successful mission like Deep Space 1 when I was at the International Space University. I am immensely thankful to my mentor at that time, Jim Taylor for sharing his considerable knowledge and for spending endless hours explaining all these exciting topics about communicating with a deep space probe. It was an amazing experience to work with the Deep Space 1 team and to participate in the tracking activities at the Mission Support Area. I am also very grateful to Dr. Kar-Ming Cheung for being a wonderful mentor, supervisor, and a brilliant mind that I have been very fortunate to work with. Thanks to that work, I coauthored the Deep Space Communications and Navigation Systems Center of Excellence report on Deep Space 1 Telecommunications. Thanks to Dr. Vilnrotter for taking me to Goldstone to complete my tour of the three DSN complexes.

Special thanks to the Zontian International Foundation for the two Amelia Earhart Fellowship Awards, and also thanks to the Zontians of the L.A. area, especially Dr. Sharon Langenbeck, Inez López, and Tricia Vick for their motivation and encouragement.

I also need to thank the International Space University in Strasbourg (France), the European Space Agency (ESA), and the National Aerospace Institute (I.N.T.A.) in Madrid (Spain), for giving me the first opportunities to work on space research.

Furthermore, I consider myself very fortunate for sharing these years with so many brilliant individuals in such a unique environment at Caltech and JPL. I would like to thank all my friends and colleagues who through the years have enriched my life through their intellect, advice, humour, and generosity, especially, Gabriela Alexandru, Jim Burke, Matthieu Liger, Reece Lumsden, Erika Podest, Carlos Romero, Myriam Ruíz, and Ian Ruíz, and all my colleagues of the Communications Group at Caltech.

Finally, I would like to thank my family and especially my mother Alicia whom I admire and love with all my heart, and who has always encouraged me to pursue all my dreams. I have been very fortunate to grow up in a loving family that has provided me with all the education and moral values that ultimately led me to where I am today.

List of Publications

- [1] Michela Muñoz Fernández, Victor A. Vilnrotter, Ryan Mukai, and Babak Hassibi. March 2006. “Coherent optical array receiver experiment: Design, implementation and BER performance of a multichannel coherent optical receiver for PPM signals under atmospheric.” *Proc. SPIE Photonics West, Free-Space Laser Communication Technologies XVIII*, G. Stephen Mecherle, Editor, vol. 61050R.
- [2] Michela Muñoz Fernández, Babak Hassibi, Ryan Mukai, and Victor A. Vilnrotter. 2005. “Multichannel Coherent Optical Receiver for PPM Signals in the Presence of Atmospheric Turbulence.” *Submitted to IEEE Transactions on Communications*.
- [3] Michela Muñoz Fernández, Victor A. Vilnrotter, Babak Hassibi, and Ryan Mukai. 2005. “Coherent Optical Array Receiver Experiment: BER performance of a multichannel coherent optical receiver for PPM signals under atmospheric turbulence.” Submitted to *Journal of Lightwave Technology*.
- [4] M. Munoz Fernandez and V. A. Vilnrotter. 2005. “Optical System for Reception of Coherently Detected PPM Signals in the Presence of Atmospheric Turbulence.” Accepted to *Journal Acta Astronautica, International Academy of Astronautics (IAA)*.
- [5] Michela Muñoz Fernández and Victor A Vilnrotter. May 15, 2005. “Optical System for Reception of Coherently Detected PPM Signals in the Presence of Atmospheric Turbulence,” *Interplanetary Network Progress Report, IPN* vol. 42-161. Joseph H. Yuen

Editor in Chief. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.

[6] Michela Muñoz Fernández and Victor A. Vilnrotter. April 2005. “Coherent optical receiver for PPM signals under atmospheric turbulence.” *Proc. of SPIE Photonics West, Free-Space Laser Communication Technologies XVII*, vol. 5712, G. Stephen Mecherle, Editor, pp. 162-173

[7] Michela Muñoz Fernández and Victor A. Vilnrotter. March 2005. “Coherent Optical Receiver for PPM Signals under Atmospheric Turbulence.” *Proc. IEEE Aerospace Conference*, pp. 1–8.

[8] M. Muñoz Fernández and V. A. Vilnrotter. October 2004. “Optical system for reception of coherently detected PPM signals in the presence of atmospheric turbulence.” *International Astronautical Congress, Space Communications Symposium, Advanced Systems*, Vancouver, Canada.

[9] Victor A. Vilnrotter and Michela Muñoz Fernández. 2004. “Coherent Detection of High-Rate Optical PPM Signals via Focal-Plane Array Detectors.” (NTR 40974) *NASA Tech Brief* 40974.

[10] M. Muñoz Fernández and V. A. Vilnrotter. November 2004. “Adaptive combining of coherently detected PPM signals in the presence of atmospheric turbulence via focal plane array.” *Proc. SPIE Remote Sensing Europe, Optics in Atmospheric Propagation and Adaptive Systems VII*, John D. Gonglewski, Karin Stein, Editors, vol. 5572, pp. 175-186

[11] M. Muñoz Fernández and V. A. Vilnrotter. April 2004. “Performance analysis and

preliminary experimental verification of a coherent optical receiver for PPM signals in the presence of atmospheric turbulence.” *Proc. 5th International Conference on Space Optics* (ICSO 2004), Toulouse, France. Ed. by B. Warmbein. ESA SP-554, Noordwijk, Netherlands: ESA Publications Division, ISBN 92-9092-865-4, pp. 411-418.

[12] M. Muñoz Fernández and V. A. Vilnrotter. June 2004. “Coherent optical receiver for PPM signals received through atmospheric turbulence: performance analysis and preliminary experimental results.” *Proc. SPIE Photonics West, Free-Space Laser Communication Technologies XVI*, G. S. Mecherle, Cynthia Y. Young, John S. Stryjewski, Editors, vol. 5538, pp. 151-162.

[13] J. Taylor. M. Muñoz Fernández, A. Bolea Alamanac, Kar-Ming Cheung October 2001. “Deep Space 1 Telecommunications.” *Design and Performance Summary Series*, Deep Space Communications and Navigation Systems Center of Excellence (DESCANSO), Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, U.S.A., http://descanso.jpl.nasa.gov/DPSummary/All_Article2.pdf

Abstract

The performance of a coherent free-space optical communications system operating in the presence of turbulence is investigated. Maximum Likelihood Detection techniques are employed to optimally detect Pulse Position Modulated signals with a focal-plane detector array and to reconstruct the turbulence-degraded signals.

Laboratory equipment and experimental setup used to carry out these experiments at the Jet Propulsion Laboratory are described. The key components include two lasers operating at 1064 nm wavelength for use with coherent detection, a 16 element (4 X 4) InGaAs focal-plane detector array, and a data-acquisition and signal-processing assembly needed to sample and collect the data and analyze the results. The detected signals are combined using the least-mean-square (LMS) algorithm. In the first part of the experimental results we show convergence of the algorithm for experimentally obtained signal tones in the presence of atmospheric turbulence. The second part of the experimental results shows adaptive combining of experimentally obtained heterodyned pulse position modulated (PPM) signals with pulse-to-pulse coherence in the presence of simulated spatial distortions resembling atmospheric turbulence. The adaptively combined PPM signals are phased up via an LMS algorithm suitably optimized to operate with PPM in the presence of additive shot noise. A convergence analysis of the algorithm

is presented, and results with both computer-simulated and experimentally obtained PPM signals are analyzed.

The third part of the experimental results, in which the main goal of this thesis is achieved, includes an investigation of the performance of the Coherent Optical Receiver Experiment (CORE) at JPL. Bit Error Rate (BER) results are presented for single and multichannel optical receivers where quasi shot noise-limited performance is achieved under simulated turbulence conditions using noncoherent postdetection processing techniques. Theoretical BER expressions are compared with experimentally obtained BER results, and array combining gains are presented. BER results are shown as a function of signal-to-noise ratio (SNR), photons per symbol, and photons per bit (PPB).

Table of Contents

| | |
|---|--------------|
| ACKNOWLEDGEMENTS | V |
| LIST OF PUBLICATIONS | VIII |
| ABSTRACT | XI |
| TABLE OF CONTENTS | XIII |
| LIST OF FIGURES | XVI |
| LIST OF TABLES | XVIII |
| GLOSSARY OF ACRONYMS..... | XIX |
| CHAPTER 1: INTRODUCTION..... | 1 |
| 1.1 Motivation to use optical technologies for deep space communications | 1 |
| 1.2 Laser communications: Coherent versus direct optical communications | 10 |
| 1.3 Statement of the problem caused by atmospheric turbulence..... | 14 |
| 1.4 Possible solutions | 16 |
| 1.5 Outline of the thesis..... | 16 |
| 1.6 Contributions of the thesis..... | 21 |
| CHAPTER 2: DESIGN OF A COHERENT OPTICAL RECEIVER ARRAY DETECTOR..... | 22 |
| 2.1 Existing space optical communication systems..... | 22 |
| 2.2 Design methodology of an optical receiver to detect signals from deep space | 24 |
| 2.2.1 Transmitter technologies..... | 25 |
| 2.2.1.1 Transmitter source | 25 |
| 2.2.1.2 Transmitter modulation formats | 29 |
| 2.2.2 Optical channel | 32 |

| | |
|---|------------|
| 2.2.3 Receiver architecture | 32 |
| CHAPTER 3: PERFORMANCE ANALYSIS OF THE COHERENT OPTICAL RECEIVER ARRAY | 49 |
| 3.1 Introduction | 49 |
| 3.2 Probability of bit error for coherent (heterodyne and homodyne) optical receiver for pulse position modulated (PPM) signals | 53 |
| 3.3 Probability of bit error for non-coherent (heterodyne) optical receiver with random phase channels for pulse position modulated (PPM) signals..... | 61 |
| 3.3.1 Case I: Probability of bit error for the single channel..... | 61 |
| 3.3.2 Case II: Probability of bit error with array combining..... | 69 |
| 3.4 Probability of bit error for non-coherent (heterodyne) optical receiver with random phase channels for Pulse Position Modulated (PPM) signals in the presence of leakage in the noise slots.... | 69 |
| 3.4.1 Expressions specific to the case with leakage | 75 |
| 3.4.2 Obtaining the SNR, signal, noise, and leakage parameters..... | 76 |
| CHAPTER 4: RESULTS OF THE COHERENT OPTICAL RECEIVER EXPERIMENT (CORE) | 83 |
| 4.1 Description of the experimental setup | 83 |
| 4.2 Adaptive combining of beatnotes using the LMS algorithm | 90 |
| 4.2.1 Adaptive combining simulated data: Signal tone and 32-PPM signals..... | 91 |
| 4.2.2 Experimental results I: Convergence of LMS algorithm for unmodulated beatnote under ideal conditions and with atmospheric turbulence | 95 |
| 4.2.3 Experimental results II: Convergence of the LMS algorithm with PPM signals and no atmospheric turbulence..... | 107 |
| 4.2.4 Convergence of LMS algorithm in the presence of spatial distortions caused by a static plexiglass plate in the optical path | 111 |
| CHAPTER 5: THE LMS ALGORITHM AND ITS APPLICATION TO COHERENT OPTICAL SIGNAL RECEPTION | 115 |
| 5.1 The signal model..... | 116 |
| 5.2 Gradient descent and LMS..... | 118 |
| 5.3 Non-stationarity of the input signal and convergence time | 120 |
| 5.3.1 Case I: No leakage present..... | 122 |
| 5.3.2 Case II: Leakage present..... | 131 |
| 5.4 More on the effects of leakage on the eigenstructure | 132 |
| 5.5 LMS with a constant target | 136 |

| | |
|---|-----|
| 5.6 LMS convergence with real world data..... | 139 |
|---|-----|

| | |
|---|------------|
| CHAPTER 6: INVESTIGATION OF PERFORMANCE OF THE COHERENT OPTICAL RECEIVER EXPERIMENT (CORE): COMPARISON OF EXPERIMENTAL RESULTS WITH THEORY | 152 |
|---|------------|

| | |
|---|-----|
| 6.1 Overview and methodology | 152 |
| 6.1.1 Operation of the software receiver | 155 |
| 6.1.1.1 Overview | 155 |
| 6.1.1.2 Correction for frequency offset using FFT centering | 156 |
| 6.1.1.3 Decision feedback system | 157 |
| 6.1.1.4 Decision feedback operation | 158 |
| 6.1.2 Synchronization issues in the optically coherent PPM system..... | 159 |
| 6.2 Evaluation of experimental system performance and comparison with theory..... | 166 |
| 6.2.1 BER performance bounds..... | 166 |
| 6.2.2 Results with the single detector | 169 |
| 6.2.3 Results with the focal plane array | 180 |
| 6.3 Conclusions | 195 |

| | |
|---|------------|
| CHAPTER 7: CONCLUSIONS AND FUTURE DIRECTIONS | 197 |
|---|------------|

| | |
|----------------------------------|-----|
| 7.1 Summary of key results | 197 |
| 7.2 Future directions | 199 |

| | |
|---|------------|
| APPENDIX: A MATHEMATICAL CHARACTERISTIC OF LEAKAGE | 201 |
|---|------------|

| | |
|---------------------------|------------|
| BIBLIOGRAPHY | 204 |
|---------------------------|------------|

List of Figures

| | |
|---|----|
| Figure 1. Future data return needs relative to current capabilities..... | 4 |
| Figure 2. Comparison of beamspreads of optical and RF communications from Saturn..... | 5 |
| Figure 3. Interactions among different organizations centered at the ITU for spectrum management. | 8 |
| Figure 4. Optimized cost model for reception systems in optical, Ka-band, and X-band..... | 9 |
| Figure 5. Electromagnetic radiation spectrum. | 10 |
| Figure 6. High-bit-rate space communication network based on future free-space laser communications. | 23 |
| Figure 7. Evolution of the demonstration experiments in space laser communications from 1994 to 2004. | 24 |
| Figure 8. Diagram of an optical communication system..... | 25 |
| Figure 9. Generic transimpedance front-end receiver configuration. | 43 |
| Figure 10. Functional diagram of the 1611 photoreceiver. | 44 |
| Figure 11. Responsivity of 1611 New Focus InGaAs photodetector. | 46 |
| Figure 12. Detailed view of the 4 X 4 FPA of InGaAs detectors..... | 48 |
| Figure 13. Diagram of the Coherent Optical Receiver Experiment (CORE)..... | 49 |
| Figure 14. Pulse position modulation format. | 53 |
| Figure 15. BER for homodyne detection. | 60 |
| Figure 16. BER for heterodyne detection..... | 61 |
| Figure 17. Coherently detected PPM beatnotes in the presence of atmospheric turbulence channels show leakage signal in noise slots example 1. | 81 |
| Figure 18. Coherently detected PPM beatnotes in the presence of atmospheric turbulence channels show leakage signal in noise slots example 2. | 82 |
| Figure 19. Coherently detected PPM beatnotes in the presence of atmospheric turbulence channels show leakage signal in noise slots example 3. | 82 |
| Figure 20. Coherent combining experiment at the Jet Propulsion Laboratory, NASA. | 84 |
| Figure 21. Snapshot of an individual PPM pulse beatnote. | 86 |
| Figure 22. Intensity distribution of the signal beam under ideal conditions. | 87 |
| Figure 23. Beam profile under spatial distortions resembling atmospheric turbulence generated with a plexiglass plate in the laboratory..... | 87 |
| Figure 24. Beam profile of the signal laser (left) and local oscillator (right) beams on the detector surface..... | 88 |
| Figure 25. Perfect alignment of signal and local oscillator beams at the InGaAs detector surface. | 88 |
| Figure 26. Sampled sequences of 4 channels containing PPM-modulated 6 MHz beatnotes under ideal conditions..... | 89 |
| Figure 27. Sampled sequences of 4 channels containing PPM-modulated 6 MHz beatnotes in the presence of atmospheric turbulence..... | 90 |
| Figure 28. Comparison of output convergence for signal tone and 32-PPM signal (real part of LMS output) | 93 |
| Figure 29. Comparison of convergence for signal tone and a 32-PPM signal with a 16-detector array $\mu = 0.003$ | 94 |
| Figure 30. Comparison of convergence for signal tone and a 32-PPM signal with a 16-detector array with $\mu = 0.008$ | 95 |
| Figure 31. Combined output signal with $\mu = 10$ | 96 |
| Figure 32. Phase of the weights with $\mu = 10$ | 97 |
| Figure 33. Error signal for $\mu = 10$ | 97 |

| | |
|--|-----|
| Figure 34. Combined output power with $\mu = 100$. | 98 |
| Figure 35. Phase of the weights with $\mu = 100$. | 99 |
| Figure 36. Error signal with $\mu = 100$. | 100 |
| Figure 37. Combined output signal with $\mu = 1000$. | 101 |
| Figure 38. Phase of the weights of the four different channels with $\mu = 1000$. | 102 |
| Figure 39. Error signal with $\mu = 1000$. | 103 |
| Figure 40. Combined output power of the beatnote signal in the presence of simulated atmospheric turbulence with $\mu = 20$. | 104 |
| Figure 41. Phase of the weights with $\mu = 20,000$. | 105 |
| Figure 42. Error signal with $\mu = 20,000$. | 106 |
| Figure 43. Combined output with $\mu = 1$. | 108 |
| Figure 44. Combined output with $\mu = 7$. | 109 |
| Figure 45. Phase of the weights for $\mu = 7$. | 110 |
| Figure 46. Combined output and weighted signal components with $\mu = 7$. | 111 |
| Figure 47. Combined output with $\mu = 8$. | 112 |
| Figure 48. Combined output with $\mu = 22$. | 112 |
| Figure 49. Phases of the weights for $\mu = 22$. | 113 |
| Figure 50. Combined output with $\mu = 22$. | 113 |
| Figure 51. Magnitude of the reference signal (blue) and the LMS output (red). The LMS output shows an apparent delay because it makes one or two samples for the LMS algorithm to converge, as predicted by Eq. (5.3.22) for $N = 4$. | 129 |
| Figure 52. LMS convergence with decision feedback and with $\mu = 6.31$. | 141 |
| Figure 53. LMS instability with $\mu = 7.94$ and decision feedback. | 142 |
| Figure 54. The behavior of LMS with $\mu = 10$ but with a perfect training sequence. | 143 |
| Figure 55. LMS Convergence with $\mu = 100$ but with a perfect training sequence. | 144 |
| Figure 56. Convergence with $\mu = 1000$ and with a perfect target sequence. | 145 |
| Figure 57. LMS loss of stability at $\mu = 10,000$. Perfect training sequences are of limited value here. | 146 |
| Figure 58. Tracking of the input phase of channel 1 by the first weight. | 148 |
| Figure 59. Tracking of the input phase of channel 2. | 149 |
| Figure 60. Tracking of the input phase of channel 3. | 150 |
| Figure 61. Tracking of the input phase of channel 4. | 151 |
| Figure 62. Block diagram of the software receiver. | 156 |
| Figure 63. Example of time misalignment prior to correction. There are five FPA channel signals that suffer a delay of approximately 10 samples relative to the modulator reference pulse. | 161 |
| Figure 64. Time misalignment between modulator reference pulses and receiver prior to correction. The misalignment only changes slowly over time. | 162 |
| Figure 65. Time-corrected pulse alignment, corresponding to the case shown in Figure 63. | 165 |
| Figure 66. Time corrected pulse alignment, corresponding to the case shown in Figure 64. | 166 |
| Figure 67. BER vs. photons per pulse with the single detector and 256 PPM. | 171 |
| Figure 68. BER vs. photons per bit with the single detector and 256 PPM. | 172 |
| Figure 69. BER vs. E_s/N_0 in dB for the single detector and 256 PPM. | 173 |
| Figure 70. BER vs. K_s assuming a hypothetical shot-noise limited single detector with 256 PPM. | 175 |
| Figure 71. BER vs. photons per bit assuming a hypothetical shot-noise limited single detector with 256 PPM. | 176 |
| Figure 72. BER vs. K_s with the single detector and 32 PPM. | 177 |
| Figure 73. BER vs. photons per bit with the single detector and 32 PPM. | 177 |
| Figure 74. BER vs. E_s/N_0 in dB for 32 PPM with the single detector. | 178 |
| Figure 75. BER vs. K_s assuming a shot-noise-limited single detector and 32 PPM. | 179 |
| Figure 76. BER vs. photons per bit assuming a shot-noise-limited detector with 32 PPM. | 180 |
| Figure 77. BER vs. E_s/N_0 with the FPA for 256 PPM. | 181 |

| | |
|---|-----|
| Figure 78. BER vs K_s for a shot noise-limited FPA system with 256 PPM..... | 183 |
| Figure 79. BER vs. photons per bit for a shot-noise-limited FPA system with 256 PPM..... | 184 |
| Figure 80. BER vs. K_s assuming an FPA with the same characteristics as the single detector (256 PPM). | 185 |
| Figure 81. BER vs. photons per bit assuming an FPA with the same characteristics as the single detector (256 PPM). | 186 |
| Figure 82. BER vs. K_s for 256 PPM with the FPA..... | 187 |
| Figure 83. BER vs. photons per bit for 256 PPM with the FPA..... | 188 |
| Figure 84. BER vs. K_s for a hypothetical shot-noise-limited FPA (32 PPM)..... | 190 |
| Figure 85. BER vs. photons per bit for a hypothetical shot-noise-limited FPA (32 PPM)..... | 190 |
| Figure 86. BER vs. K_s for a hypothetical FPA with the same noise characteristics as the single detector (32 PPM). | 191 |
| Figure 87. BER vs photons per bit for a hypothetical FPA with the same noise characteristics as the single detector (32 PPM). | 192 |
| Figure 88. BER vs K_s actually achieved with the FPA (32 PPM)..... | 193 |
| Figure 89. BER vs photons per bit actually achieved with the FPA (32 PPM). | 193 |
| Figure 90. BER vs. E_s/N_0 for 256 PPM: Illustration of single channel performance vs. combined channel. Single channel BER is plotted vs. single channel SNR and not against combined channel SNR as in previous figures. | 194 |
| Figure A1. Magnitude and phase of the first channel plotted as a function of time..... | 201 |
| Figure A2. Magnitude and phase of the second channel plotted as a function of time. The phase behavior of this channel differs greatly from that of the first..... | 202 |

List of Tables

| | |
|--|----|
| Table 1. ITU frequency allocation for deep space research..... | 6 |
| Table 2. Possible technologies for free-space systems. | 26 |

Glossary of Acronyms

| | |
|---------------|--|
| APD | Avalanche Photodiode |
| BER | Bit Error Rate |
| COSPAR | Committee of Space Research |
| DSN | Deep Space Network |
| ESA | European Space Agency |
| FO | Fiber-Optic |
| FOV | Field of View |
| FPA | Focal Plane Array |
| IAU | International Astronomical Union |
| IF | Intermediate Frequency |
| ITU | International Telecommunications Union |
| IUCAF | Scientific Committee on Frequency Allocations for Radio Astronomy and Space Science |
| JPL | Jet Propulsion Laboratory |
| LMS | Least Mean Square |
| LO | Local Oscillator |
| LWIR | Far or Long Wave IR |
| MGS | Mars Global Surveyor |
| MSA | Mission Support Area |
| NASA | National Aeronautics and Space Administration |
| NIR | Near Infrared |
| NUV | Near Ultraviolet |
| PD | Photodiode |
| PLL | phase-locked loop |
| PMT | Photomultiplier Tube |
| PPM | Pulse Position Modulation |
| RF | Radiofrequency |
| SER | Symbol Error Rate |
| SFCG | Space Frequency Coordination Group |
| SNR | Signal to Noise Ratio |
| SWIR | Shortwave IR |
| URSI | International Union of Radio Science |
| WRC | World Radiocommunications Conference |

Chapter 1: Introduction

1.1 Motivation to use optical technologies for deep space communications

For future interplanetary missions, there is a need for higher capability and smaller and more efficient space-borne telecommunications, but there are constraints due to limits on RF spectrum allocation. Optical space communications systems are becoming more practical as technology develops and offer significant advantages over radio frequency communications. The main advantages are the ability to concentrate power in narrower beams (small beam divergence), the absence of bandwidth limitations in the optical communications frequency range, and the drastic reduction in component sizes. Optical wavelengths are very short and correspond to very high carrier frequencies. Increasing the carrier frequency theoretically increases the available transmission bandwidth and, therefore, the information capacity of the system. Optical communications operate at frequencies much higher than those of RF communications. There are no bandwidth limitations imposed on optical communications at this time. It is not unusual for an optical crosslink to achieve antenna (telescope) gains in excess of 110 dBi (dB relative to an isotropic radiator) with apertures of 20 cm. This high gain translates directly into a significant reduction in the required transmitted power. Because of their highly

directional characteristics, lasers provide higher security and greater resistance to interference (low probability of intercept, and freedom from jamming).

Because the strength of a signal decreases as the square of the distance it might travel, deep space communications are extremely challenging. Near-earth modern communications satellites use geosynchronous orbits, Molniya orbits, or low earth orbits. Comparing the case of a spacecraft in the geosynchronous orbit (Arthur C. Clark, 1945, building on work by Konstantin Tsiolkovsky and on the 1929 work by Herman Potočník) transmitting at 10 Gbps vs. a deep space system at Pluto/Neptune at distances 10^5 times greater, only 1 bit/sec could be achieved due to the fact that the beam would spread over an area 10^{10} times larger.

In order to overcome spreading effects due to propagation over large distances, communication systems at the transmitter and receiver sides have been improved over the years. An example is that current links to Mars achieve data rates of the order of 10 to 100 kbps as in the case of the Mars Exploration Rovers (Taylor *et al.*, 2005). Current RF technology has already reached an extremely high level of performance since spacecraft transmitter antennas are already large and are transmitting at a high power that should not be increased due to the fact that it is very difficult to generate electrical power at long solar distances, and waste heat due to inefficiencies of the energy conversion components of the transmitter should be eliminated. On the receiver side, NASA's Deep Space Network uses 34 m and 70 m antennas in the three complexes around the globe (Goldstone, Madrid, and Canberra) and ESA uses its own 35 m antennas at New Norcia,

Australia, and Cebreros, Spain to track near-earth satellites and deep space probes. Development and maintenance costs of these ground antennas are substantial.

Instrument data volumes from space missions are increasing, but the data return capabilities for systems in deep space are orders of magnitude below what is attainable at near-Earth distances. Figure 1 (Lesh, J., 2005) shows the future data return needs relative to current capabilities for the Mars Global Surveyor (MGS) mission that has been mapping features of the Martian terrain. The horizontal axis represents data rate and the vertical blue lines on the left side represent the current MGS capability scaled to Saturn distance. For the vertical dimension, things above the central data rate axis arrow need to be further investigated while those below are thought to encourage public interest in space exploration. The oval elements refer to horizontal data rate regions where instruments are expected to operate and the regions of improvement of several communications technologies are shown. The winning technologies will be selected based on research and technology planning and on life-cycle-cost analyses. The improvements that optical communications have to offer are tremendous.

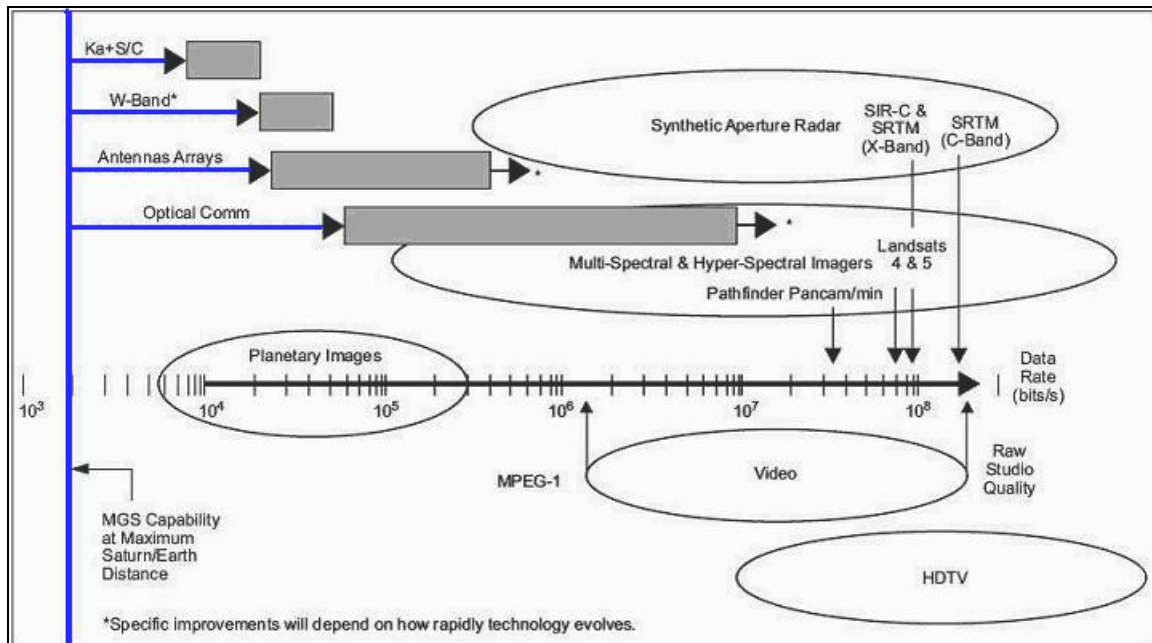


Figure 1. Future data return needs relative to current capabilities.

The DSN's RF performance has undergone an improvement of twelve orders of magnitude over the years, mainly because of a gradual increase in the carrier frequency. The DSN uses X-band, and more recently Ka-band as in the case of the Deep Space One mission that successfully tested Ka-band downlink capabilities (Taylor *et al.*, 2001). Ka-band holds a potential four-fold increase in data rate compared to X-Band. Theoretically, the improvement in going from X-band to Ka-band is 11.6 dB due to frequency squared gain, even if experimentally only 6 dB of gain have been achieved. This fact is obviously important as at the end it means reduced project cost because a higher data rate requires fewer ground resources and less mission operation support. Another advantage of using Ka-band is the availability of greater bandwidth as NASA and other agencies move away from frequency bands that face increasing encroachment by personal communications

systems and other emerging information technologies. Following the same trend, if even higher frequencies of approximately 3×10^5 GHz are used, as in the case of optical communications, the improvement is potentially immense.

Illustrating a comparison of the relative beam spreads of RF and optical frequencies, Figure 2 (Lesh, 2005) shows that for an optical wavelength of $1 \mu\text{m}$, the spot size of a beam transmitted from a 10 cm optical telescope on the Voyager spacecraft is equal to one Earth diameter. If X-band had been used instead with a 3.7 m transmitter antenna, the spot size would have been equivalent to a thousand Earth diameters. Theoretically the improvement is 90 dB, although experimentally 60 dB could be achieved with current technology.

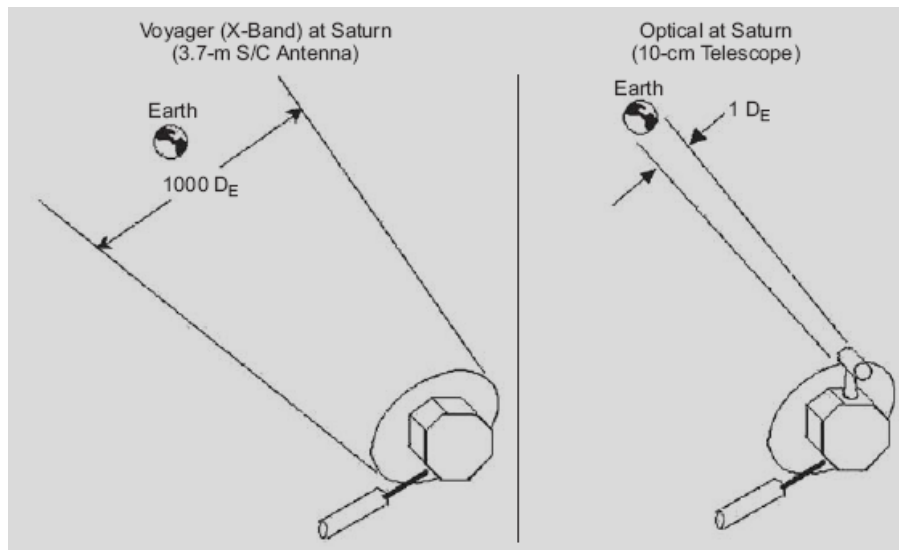


Figure 2. Comparison of beamspreads of optical and RF communications from Saturn.

Another important point in dealing with optical frequencies is the fact that their use has not been regulated yet. The International Telecommunications Union (ITU) is

responsible for coordinating spectrum management at the global level. The ITU allocates the frequency ranges for use in deep space research (for spacecraft greater than 2 million km from Earth) as shown in Table 1:

| Frequency Band | Carrier Frequency Range |
|------------------|-------------------------|
| S-band uplink | 2110-2120 MHz |
| S-band downlink | 2290-2300 MHz |
| X-band uplink | 7145-7190 MHz |
| X-band downlink | 8400-8440 MHz |
| Ka-band uplink | 34,200-34,700 MHz |
| Ka-band downlink | 31,800-32,300 MHz |

Table 1. ITU frequency allocation for deep space research.

The global framework for spectrum management is provided by the Radio Regulations of the ITU, which have international treaty status and thus are binding for all members of the ITU (Van Driel, 2004). They provide rules to national administrations that allow them to regulate equitable access to the radio spectrum for all entities requiring frequency allocations. The Radio Regulations contain the international Frequency Allocation Table. In order to modify the Radio Regulations, the ITU organizes a World Radiocommunication Conference (WRC) once every three years on average, which is attended by over 2000 representatives of 180+ national administrations and other accredited organizations, such as IUCAF (the Scientific Committee on Frequency

Allocations for Radio Astronomy and Space Science) and which lasts for a month. In the past, the ITU concerned itself only with regulating spectrum use at radio frequencies up to 275 GHz (1 mm wavelength), but recent developments indicate that the role of the ITU as a global body for the coordination of spectrum management will be extended over the entire electromagnetic spectrum as required. It is foreseen that at the WRC in 2010 frequency allocations will be considered between 275 GHz and 3000 GHz (100 μ m wavelength). It seems clear that the inclusion of infrared and optical wavelengths in the ITU Radio Regulations is still remote, but the process starts with providing information on optical systems and the characteristics that can be used to regulate their frequency protection. Concerning the protection of both Earth-based astronomical observations and spaced-based Earth exploration observations from intersatellite communication lasers, studies were made in the period 2002-2003, which resulted in an ITU Recommendation on this issue. Since systems operating at these unregulated frequencies cannot yet be registered at the ITU, it was decided that a register of such laser systems be kept by the Space Frequency Coordination Group (SFCG), which gathers national and international space agencies and where the IUCAF has observer status, with IUCAF as liaison to the IAU (International Astronomical Union) on the proliferation of these systems. The complex structure centered on the ITU is sketched in Figure 3. In the left half of the picture one finds different organizations dealing with astronomical research, such as COSPAR, the IAU and URS, under ICSU, and IUCAF. The right half indicates the various ITU forums where studies related to both Radio Regulations and updates of ITU Recommendations are performed.

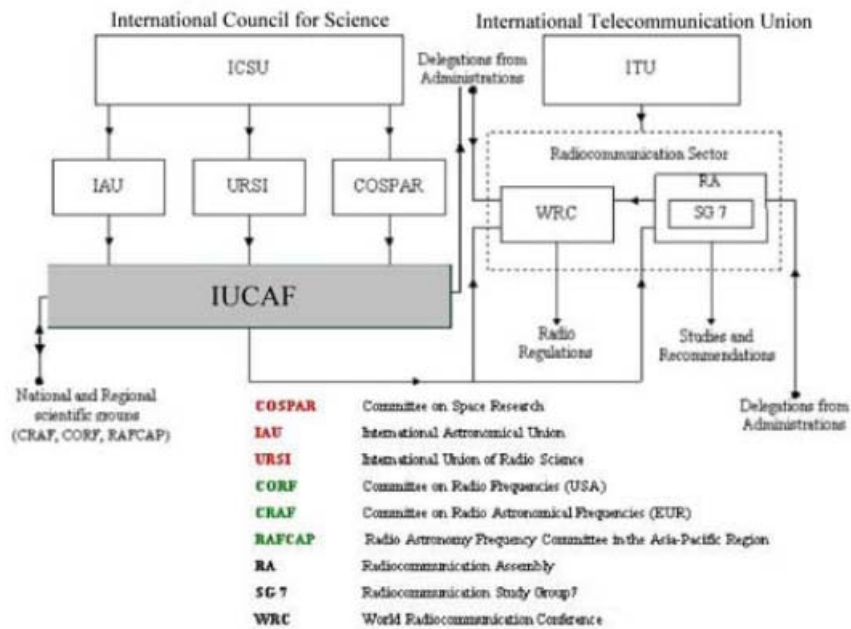


Figure 3. Interactions among different organizations centered at the ITU for spectrum management.

Research that has been conducted on optical communications will contribute to all future missions contemplating the use of optical frequencies for high-rate or long distance communications, including Mars data return, Europa Orbiter, and Pluto Flyby, enabling the use of near real-time video from Mars rovers and future human explorers. The technology developed in this effort will contribute to successful links from deep space to an optical ground station to be used with all future NASA spacecraft incorporating laser communications capabilities. However, optical technologies will not replace microwave systems for all applications because low-rate systems and applications requiring high availability from Earth are often better served by RF systems.

The high volume of data from a spaceborne network would need to be transmitted to Earth by means of ground station site diversity (Alexander and Stephen, 1997). The

optical ground stations should be far enough apart to be located in uncorrelated weather systems to ensure a high probability that at least one of the sites would have a clear view of the satellite. The ground sites would be connected using high data-rate optical fiber technology (Chapman and Fitzmaurice, 1991). JPL (Lesh and Robinson, 1986) has conducted cost studies for optical and X-band technologies. For X-band three DSN stations would be needed, while for optical technologies, nine stations would be needed. Figure 4 shows an optimized cost model for reception systems operating in Ka-band, X-band and optical frequencies representing total NASA costs (\$B) as a function of communications growth (dB), (Lesh, 2005).

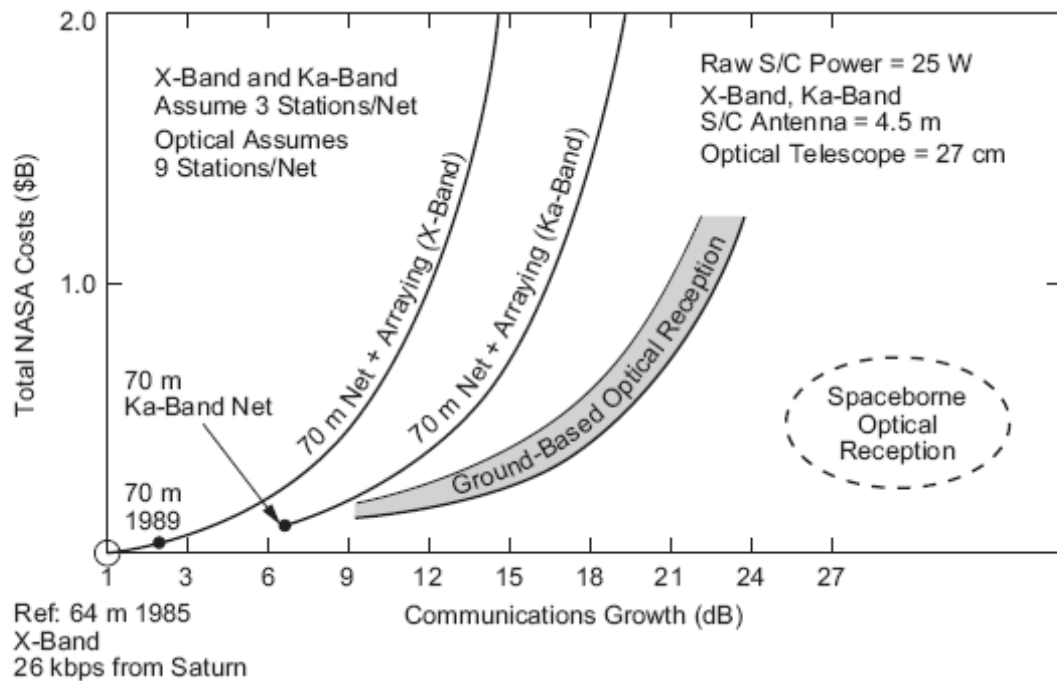


Figure 4. Optimized cost model for reception systems in optical, Ka-band, and X-band.

1.2 Laser communications: Coherent versus direct optical communications

Laser communication refers to the use of light produced by lasers in the Near Ultraviolet (NUV), Visible (Vis), Near Infrared (NIR), Shortwave IR (SWIR), and Far or Long Wave IR (LWIR) infrared portions of the electromagnetic spectrum to communicate information. The communication process in those frequency bands differs significantly from that of Radio Frequency (RF) or Fiber-Optic (FO) systems. Figure 5 shows the electromagnetic radiation spectrum.

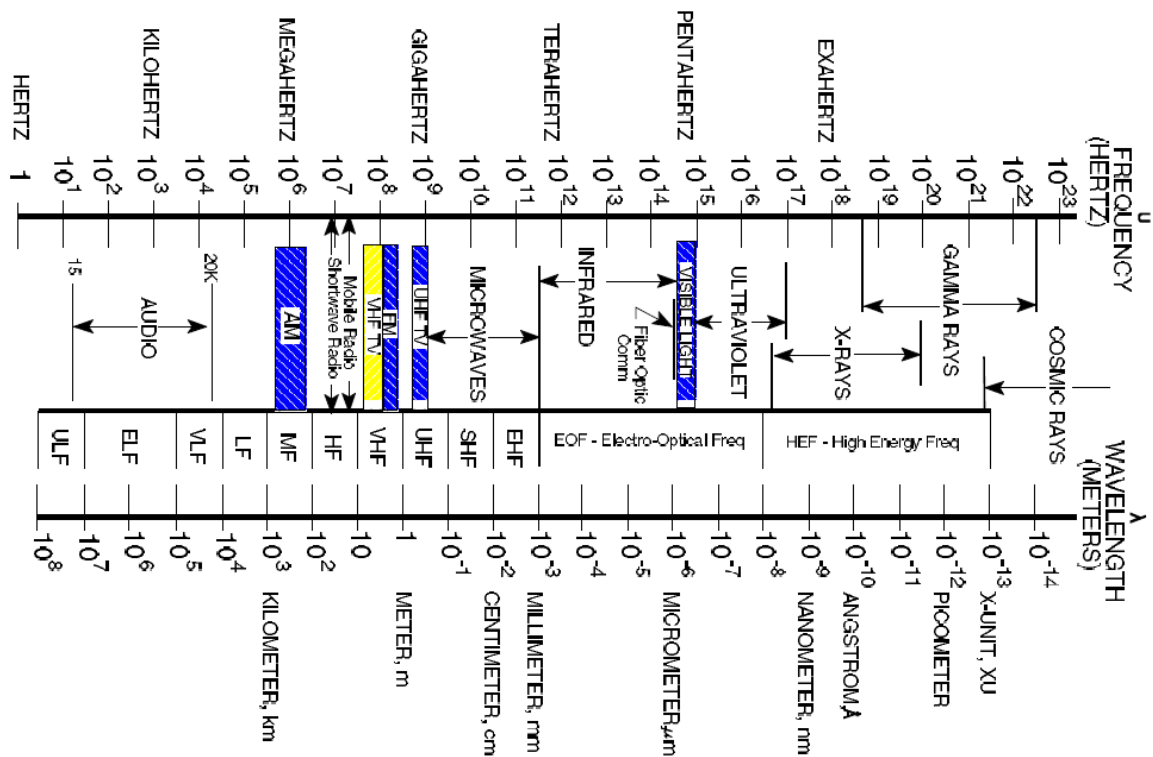


Figure 5. Electromagnetic radiation spectrum.

The performance of an optical link depends on the receiver sensitivity measured in terms of received photons per bit. It is very important to improve receiver sensitivity in order to reduce the high cost associated with increasing the transmit power and receiver aperture.

Optical receivers can be divided into two basic types (Gagliardi and Karp, 1995): power detecting receivers and heterodyne receivers. Power detecting receivers are usually called direct detection, or noncoherent, receivers. These receivers represent the simplest type for implementation and can be used whenever the transmitted information occurs in the power variation of the received field. On the other hand, in the case of heterodyne receivers (also called coherent receivers), a locally generated field is optically mixed with the received field through a front end mirror, and the combined wave is photodetected. Direct detection is more commonly used for higher frequencies (SWIR and higher), while for LWIR, coherent or heterodyne detection is used. The potential for coherent systems at shorter wavelengths exists, but such systems pose significant implementation challenges.

In digital communications (Proakis, 2000), the term “coherent” refers to those systems that employ carrier phase recovery. In coherent optical communications systems, the term “coherent” is defined differently. An optical communication system is called “coherent” as long as there is optical signal mixing, and carrier phase recovery need not be present. Even if the demodulator does not use carrier phase recovery but instead uses non-coherent or envelope detection, the optical communications system is called a coherent optical communication system due to the presence of optical mixing.

For example, in RF communications, a differential phase-shift keying (DPSK) system is considered a noncoherent communication system (Proakis, 2000), whereas in optical communication, it is considered a coherent system (Betti *et al.*, 1995, Hooijmans, 1994, Okoshi and Kikuchi, 1988, Ryu, 1995). Following the traditional terminology of coherent optical communications, coherent optical receivers with and without phase tracking would be called synchronous and asynchronous receivers, respectively. Asynchronous receivers usually employ power or envelope detection.

In terms of sensitivity, phase-modulated coherent optical communications systems generally provide the best performance with all types of modulation schemes. JPL decided to use Pulse Position Modulation for optical deep space communications for reasons that will be further explained.

The mixing of two lasers for communication purposes was considered in the earliest days of optical communications (Goodwin, 1967, Oliver, 1961). Early systems operated in free space used high power long-wavelength laser sources (DeLange, 1972, Goodwin, 1967, Nussmeier *et al.*, 1974, Peyton *et al.*, 1972). Even today, coherent space communications still has advantages over on-off keying (Chan, 1987, 2000, 2003, Rochat *et al.*, 2001), especially for inter-satellite communications. Coherent optical communication is also used for ultra dense radio-on-fiber signal (Kikuchi and Katoh, 2002a,b, Kuri and Kitayama, 2002, 2003).

High performance quantum-limited operation of optical detector arrays can be achieved with current technology in the two different ways previously mentioned, namely by the use of photon-counting detector arrays (Vilnrotter *et al.*, 2002) and by the use of

optical mixing prior to detection. It has been determined that focal plane arrays that sample the spatially degraded signal fields, together with high-speed signal processing to implement the real-time combining algorithms, provide the necessary capability for optimizing receiver performance when direct detection is employed. This technique has been evaluated through analysis and laboratory simulation with high-gain photon-counting PMT arrays and shown to provide approximately 3-5 dB gain over a large area single-detector PMT receiver covering the same FOV when operating in turbulent and high-background environments (Vilnrotter *et al.*, 2002). However, currently available PMT arrays have low quantum efficiency (less than 10%) at the wavelengths of interest (1.064 micrometers). Similar gains have been demonstrated analytically with higher quantum-efficiency (40%) APD arrays over large area single-detector APD receivers, but the performance of APD receivers is impeded by thermal noise due to their limited gain.

Coherent detection can be accomplished with detector arrays without internal gain, but requires the addition of a precisely aligned optical field. This detection scheme is not sensitive to background radiation and, hence, suffers no measurable degradation from background light under most conditions of interest. However, the detector outputs must be phased up and added in real time. Coherent detection together with focal-plane arrays provides very high gain through the coherent addition of a strong optical field to the received signal fields, producing a “signal-local” cross term after detection that overcomes thermal noise and achieves shot noise-limited performance. An additional advantage of coherent detection is that high quantum efficiency InGaAs detector arrays can be used. Since a local optical field has to be added to the received field prior to

detection, this technique is more difficult to implement, but not prohibitively so. It was initially demonstrated (Muñoz Fernández *et al.*, 2004) in the laboratory that simply “flooding” the detector array with the local field is sufficient to produce the required cross-terms, generating an easily detectable difference frequency tone at the output of every detector element. Further work has demonstrated shot-noise limited coherent detection of pulse-modulated fields with an array and determined the potential gain through the use of this technique (Muñoz Fernández *et al.*, 2005a,b,c).

1.3 Statement of the problem caused by atmospheric turbulence

Problems arise when using optical communications due to short optical wavelengths. The three main atmospheric processes that affect optical wave propagation are absorption, scattering, and refractive-index fluctuations. Index of refraction fluctuations lead to irradiance fluctuations, beam broadening, and loss of spatial coherence of the optical wave at the receiver. Absorption by water vapor reduces the energy content in the communication beam, and turbulence increases the beam's divergence. Terrestrial reception suffers from clear-air turbulence and cloud absorption effects that do not plague longer-wavelength RF systems. Here, only the clear-air turbulence problem is addressed.

The use of a laser beam as a carrier for a satellite-to-ground link enables transmission using very narrow beam divergence angles. Inhomogeneity in the temperature and pressure of the atmosphere leads to variations of the refractive index and the transmission path. Since the index of refraction of air is not uniform, electromagnetic waves passing through it will be distorted. As a result, laser communications

performance is affected by the atmosphere since it is a dynamic and imperfect medium. Atmospheric channel effects include fluctuations in the signal amplitude and phase as well as attenuation.

Therefore, a laser beam traversing the atmosphere is constantly refracted (bent), resulting in scintillation. This turbulence-induced fading impairs free-space optical links in much the same way that flat multipath fading impairs radio-frequency wireless links. These variations of refracted index as well as pointing vibrations can cause fluctuations in the intensity and phase of the received signal leading to an increase in link error probability.

In the context of optical communications, this randomization of the optical phase-front often requires the use of a larger receiver field of view, thus admitting more unwanted background radiation into the receiver. In the presence of background radiation, performance of direct detection optical receivers often degrades significantly. One way to overcome the effects of background radiation is to use coherent detection, which is generally much less sensitive to background effects than direct detection (Gagliardi and Karp, 1995). In addition, detectors used for coherent detection have higher quantum efficiency than those used for direct detection photon-counting applications.

In order to improve the optical link performance one of the important tasks is to develop new communications techniques that will mitigate the effects of atmospheric turbulence.

1.4 Possible solutions

The solution proposed here is to use focal-plane arrays to collect the optical signals from different spatial modes of the received signal field simultaneously, and then recombine the signals coherently using adaptive algorithms. Analysis and proof-of-concept demonstration of coherent adaptive array detection with PPM signals will be described in the following sections.

1.5 Outline of the thesis

A method of coherent detection of high-rate pulse-position modulation (PPM) on a received laser beam has been conceived as a means of reducing the deleterious effects of noise and atmospheric turbulence in free-space optical communication using focal-plane detector array technologies. In comparison with a receiver based on direct detection of the intensity modulation of a PPM signal, a receiver based on the present method of coherent detection performs well at much higher background levels.

In principle, the coherent-detection receiver can exhibit quantum-limited performance despite atmospheric turbulence. The key components of such a receiver include standard receiver optics, a laser that serves as a local oscillator, a focal-plane array of photodetectors, and a signal processing and data acquisition assembly needed to sample the focal-plane fields and reconstruct the pulsed signal prior to detection. The received PPM-modulated laser beam and the local-oscillator beam are focused onto the photodetector array, where they are mixed in the detection process. The two lasers are of the same or nearly the same frequency. If the two lasers are of different frequencies, then

the coherent detection process is characterized as heterodyne and, using traditional heterodyne-detection terminology, the difference between the two laser frequencies is denoted the intermediate frequency (IF). If the two signals are of the same frequency and remain aligned in phase, then the coherent detection process is characterized as homodyne (essentially, heterodyne detection at zero IF).

As a result of the inherent squaring operation of each photodetector, the output current includes an IF component that contains the signal modulation. The amplitude of the IF component is proportional to the product of the local oscillator signal amplitude and the PPM signal amplitude. Hence, by using a sufficiently strong local-oscillator signal, one can make the PPM-modulated IF signal strong enough to overcome thermal noise in the receiver circuits: this is what makes it possible to achieve near quantum-limited detection in the presence of strong background.

Following quantum-limited coherent detection, the outputs of the individual photodetectors are automatically aligned in phase by use of one or more adaptive array compensation algorithms (e.g., the least-mean-square (LMS) algorithm). Then the outputs are combined, and the resulting signal is processed to extract the high-rate information as though the PPM signal were received by a single photodetector.

The thesis has been structured following the subsequent order. Chapter 1 will describe the role of optical communications in deep-space exploration and high data-rate terrestrial communications: inherently greater energy concentration than comparable cost and weight RF systems; greater useful bandwidth than conventional RF systems; and order-of-magnitude smaller required transmitter and receiver apertures than comparable

X-band or Ka-band systems. Problems encountered due to short optical wavelengths include clear-air turbulence and cloud absorption effects that do not plague longer-wavelength RF systems. Here, we address only the clear-air turbulence problem.

Chapter 2 will focus on the design of a coherent optical communications system concept for turbulence mitigation. In the absence of background interference, the capacity of unconstrained optical pulse-position modulation (PPM) for the photon counting channel is infinite (Pierce, 1978). This inherently high photon information-efficiency prompted consideration of PPM for deep-space communications. PPM modulation at the transmitter could be either coherent or non-coherent from pulse to pulse.

Chapter 3 will focus on optical focal-plane array detection theory and on the coherent optical receiver for turbulence-degraded signals. The theoretical model consists of a lens (representing the receiving telescope); a local laser field mixed with the received signal field; a focal-plane detector array for optically detecting the sum fields; optical detection options (homodyne, heterodyne); an array of amplifiers providing suitable signal levels to the adaptive combining algorithms; adaptive algorithms for optimally weighing and combining electrical array signals designed specifically for PPM; and combined-signal PPM symbol detection options (coherent homodyne and heterodyne, non-coherent baseband). In modal homodyne detection of turbulence-degraded optical fields, the subjects to be covered include the local optical field matched to optical phase of received signal in each spatial mode; the model of array-detected homodyne signal fields and shot-noise spectral level; the comparison with single spatial-mode signal field

of same power; the SNR of homodyne-detected multi-mode array signal and equivalent single-mode signal; and the theoretical PPM symbol-error probabilities for homodyne-detected optimally combined array signal fields. In modal heterodyne detection of turbulence-degraded optical fields the subjects to be covered include: local optical field matched to array dimensions, but not to each spatial mode; model of array-detected heterodyne signal fields and shot-noise spectral level; SNR of heterodyne-detected multi-mode array signal and equivalent single-mode signal; and theoretical PPM symbol-error probabilities for heterodyne-detected optimally combined array signal fields.

Experimental results are first described in chapter 4. The description of optical FPA receiver laboratory demonstration includes: local and signal lasers, external modulator producing pulse-to-pulse coherence, turbulence simulator, lens, beam splitter, optical detector array, amplifier array, GaGe data-acquisition card; description of PPM modulator; spectral levels of near shot-noise limited single detector, and shot-noise plus thermal-noise spectral levels. Experimental data analysis includes: digital filtering of data and complex baseband downconversion; description of data quality, including difficulty of obtaining stable PPM-modulated beat-frequency at ~ 6 MHz; use of recorded PPM symbol-stream to provide a perfect reference; processing of recorded data using both a perfect reference and averaged-PPM symbols; comparison of convergence rates for recorded data processed with known and averaged PPM data; description of coherent and non-coherent detection software receiver; comparison of theoretically detected PPM SER (BER) with theoretical results, for both single detector (near shot-noise limited) and array detector (experimentally measured shot-noise plus thermal noise spectral levels). The

study of the modified LMS algorithm applied to PPM symbols includes modified LMS algorithm applied to noiseless PPM symbols: instantaneous gradient estimate via correlation of the received signal vector with the error signal for known PPM symbols; equivalence of correlation using averaged PPM symbols and known PPM symbols, in the absence of noise; characteristics of phase-aligned combined signals; and suitability of coherent and non-coherent symbol detection techniques. The effects of noise on LMS convergence and tracking, decision feedback, and performance of decision-aided averaged-PPM LMS via analysis and simulation are also included.

Chapter 5 reviews the mathematical theory behind the LMS algorithm and examines the application of LMS to coherently combined PPM signals. The convergence behavior of the LMS algorithm is analyzed in light of the time-varying eigen structure that results from the non-stationarity of the inputs. Chapter 6 investigates receiver performance in terms of BER as a function of photons per bit and signal-to-noise ratios and compares the laboratory experimental results with theory, demonstrating shot noise limited performance for the single-channel coherent receiver. Chapter 7 summarizes key results of this research and discusses possible future directions. It includes significance of results; utility of non-coherent detection for intensity-modulated PPM (lack of pulse-to-pulse coherence at the transmitter); future work including theoretical derivation of quantum-optimum receiver for turbulence-degraded optical fields; Reed-Solomon coding for coherently detected PPM to improve BER performance; and implications for future optical deep-space communications.

1.6 Contributions of the thesis

The main contributions of the thesis are:

1. First time an optical coherent (heterodyne) system with PPM modulation has been designed and implemented for deep space communications that is capable of compensating for atmospheric turbulence distortions.
2. Use of PPM modulation for coherent (heterodyne) optical communications.
3. Demonstration of shot noise limited performance for an optical coherent (heterodyne) receiver based on a focal plane array of InGaAs detectors for deep space communications applications.
4. Application of the LMS algorithm to heterodyned PPM signals to compensate for atmospheric turbulence.
5. Demonstration of signal array combining gain by means of the LMS algorithm for an optical coherent receiver.

Chapter 2: Design of a Coherent Optical Receiver

Array Detector

2.1 Existing space optical communication systems

Multigigabit communications, widely introduced in terrestrial fiber-optic communication networks, will have an important role in future space communications. Figure 6 shows a high-bit-rate space communication network based on future free-space laser communications (Arimoto, 2004).

For near-terrestrial space optical systems, fine pointing systems (1 to 10 μrad accuracy) and very high bandwidths ($\geq 10\text{Gbps}$) are needed. For deep space communication, in order to be competitive with RF communications, optical links must provide data rates on the order of 30 Mbps with extremely fine pointing accuracies of less than $1\mu\text{rad}$. Deep space communications need high sensitivity receivers, large ground receiver apertures, and minimal power and mass at the transmitting end.

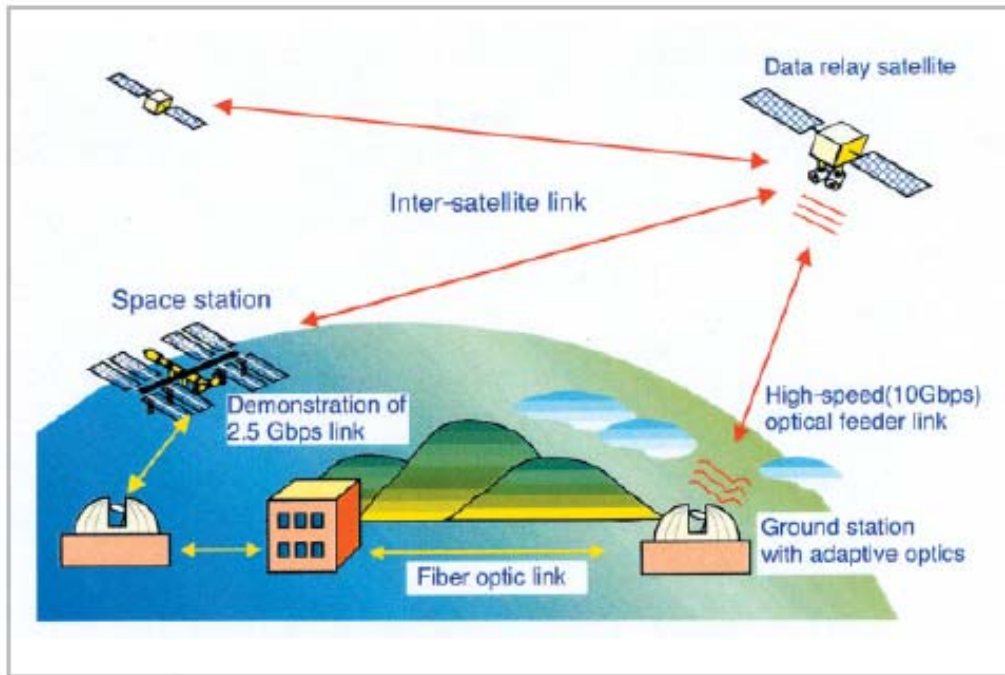


Figure 6. High-bit-rate space communication network based on future free-space laser communications.

Figure 7 shows the evolution of the demonstration experiments in laser communications from 1994 to 2004 including ETS-VI (LCE) with the collaboration of JPL, NASDA and CRL (Communications Research Laboratory, Japan) in the GOLD experiment; the flight demonstration component includes the laser communication package called the Laser Communication Experiment (LCE) on Engineering Test Satellite VI (ETS-VI), and a flight demonstration called OICETS linking a LEO satellite with the European experimental communication satellite ARTEMIS; and SILEX (Spot 4), the world's first launch-ready civilian laser communication system. The first in-orbit test was in December 2001, and the system has been operational since 2003 with Bit rates up to 50 Mbps. There are several ongoing European developments in preparation for LEO and GEO

communication terminals. The best result to date on coherent (homodyne) detection in space communications is by Donnier (European company) for intersatellite links that has demonstrated 20 photons/bit at 565 Mbps.

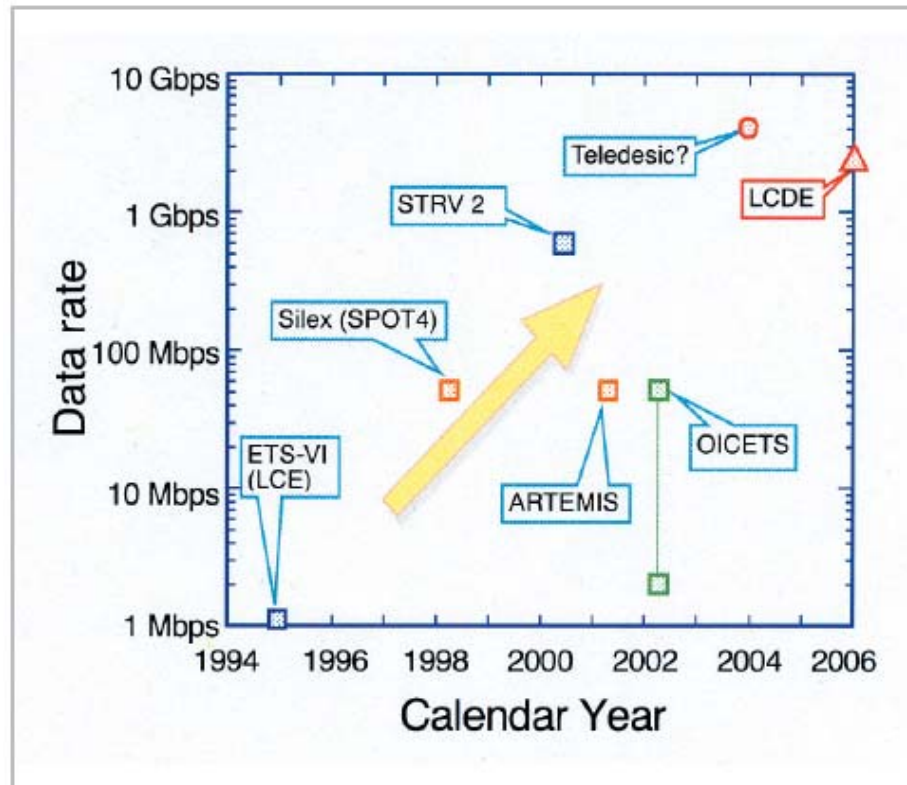


Figure 7. Evolution of the demonstration experiments in space laser communications from 1994 to 2004.

2.2 Design methodology of an optical receiver to detect signals from deep space

The goal of this thesis is to design, implement and analyze the performance of an optical receiver for deep space communications. Figure 8 shows a diagram of an optical communication system including the technologies involved (Swanson, 1994).

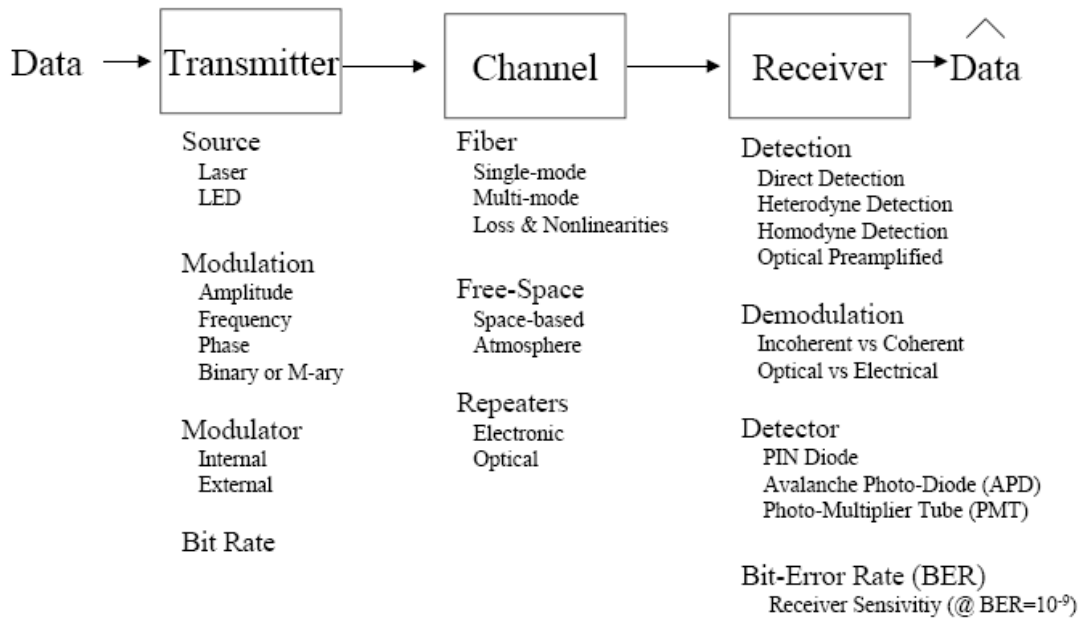


Figure 8. Diagram of an optical communication system.

2.2.1 Transmitter technologies

2.2.1.1 Transmitter source

The choice of laser technology is the starting point when designing a free-space system as it drives the operating wavelength and usually determines the modulation format to be used. This format, in turn, will determine the type of photodetection and demodulation employed at the receiver (Alexander, 1987). Table 2 describes possible laser technologies for free-space systems.

| Laser Type | Examples | Operating Wavelengths | Drawbacks |
|---------------|-------------------------|-----------------------------|---------------------------------|
| Semiconductor | AlGaAs InGaAsP | 830 nm 980 nm 1550 nm | Low powers |
| Solid state | Nd:YAG | 1060 nm | Separate modulator |
| Gas | CO ₂ HeNe | 10,600 nm 530 nm | Complexity Lifetime issues |
| Fiber | Erbium doped fiber | 1550 nm | Not tested in space environment |

Table 2. Possible technologies for free-space systems.

The CO₂ (carbon dioxide) laser and the HeNe (helium neon) gas lasers were regarded as good candidates at the beginning of the research on space optical communications (McElroy *et al.*, 1997) but as more research was conducted, it soon became evident that the 10 μ m CO₂ laser was not the winning technology for use in space because of weight, lifetime, and operational problems. It should be noticed that it also had good qualities such as high power, efficiency, narrow linewidth, broad tunability, and excellent atmospheric penetration. The linewidth of a laser, typically a single-frequency laser, is the width of the spectral power spectral density of the electric field of its output.

It is strongly related to the temporal coherence, characterized by the coherence time or coherence length.

Towards the end of the 1970s, semiconductor diode lasers operating at room temperature became available, providing a very promising transmitter source for optical intersatellite links. In 1980, ESA conducted the first studies to explore the potential of using these new devices for intersatellite links. Semiconductor lasers (GaAs, InGaAs, GaN) (Streifer, 1989, and Chan *et al.*, 1983) are very attractive due to their high efficiency, long life, high pulse repetition frequency, small size and their ability to be directly modulated by varying their bias current. However, these only provide a hundred milliwatts of usable optical power, which is not enough for deep space communications.

Regarding solid-state lasers (Users Manual, Lightwave Electronics, 1988), Nd:YAG is the best example, where one or more 0.8 μm high power AlGaAs diode arrays are used to optically pump a crystal composed of Nd:YAG that will emit at 1.06 μm . Nd:YAG is capable of providing a few watts of usable continuous-wave (CW) output power but is not very efficient. It needs a separate modulator which introduces additional optical alignment requirements and optical losses. Pulsed mode operation is also allowed using cavity-dumping or Q-switching. Coherent systems based on Nd:YAG lasers are highly promising for high-data-rate systems. There is no restriction, in principle, to the achievable laser power, and detector sensitivity can almost reach the theoretical quantum limit. Nd:YAG has continued to be useful as diode pumping has improved life, efficiency, and beam quality: it represents the best candidate for long range applications in space.

Fiber lasers are a more recent development for use in the telecommunications industry. This type of laser is typically fabricated with erbium doped fiber pumped at 980 nm or 1480 nm. They have not been considered because they have not been tested in a space environment.

For the experiments conducted in this thesis, the technology selected for the laser source has been an Nd:YAG laser at 1064 nm. It is a diode pumped solid state ring laser, series 120 by Lightwave Electronics (Users Manual, Lightwave Electronics, 1990) class IIIb laser as defined by the Federal Register 21 CFR 1040.10 Laser Safety Standard. It is a single mode continuous wave output laser that produces an ultranarrow linewidth, frequency stable, single axial mode beam. It incorporates diode pumping for reliability, compactness and efficiency. Inside the unit, there is a monolithic traveling wave ring resonator that ensures stable, single frequency operation over a wide range of operating conditions. It features coarse frequency tuning with a digital counter and a clock as well as fine frequency tuning; both have been used to adjust the desired intermediate frequency in the experiments. The beam is linearly polarized, collimated and single mode (TEM_{00}). It provides kHz short term stability and MHz long term stability. The linewidth (1ms) provided by the manufacturer is 5 kHz and jitter (1sec) is 75 kHz. As in the experiment, heterodyne detection was performed, two lasers were used: a lower power signal laser (Model 120-01 with CW output power of 4 mW) and a higher power local oscillator (Model 120-03 with CW output power of 40 mW). The extremely stable, ultra narrow linewidth output of the series 120 lasers makes them prime candidates for

coherent communications, LIDAR, and other scientific as well as commercial applications.

2.2.1.2 Transmitter modulation formats

A laser can be amplitude, frequency, phase, or polarization modulated. There is a wide variety of possible transmitter symbol sets that are used in optical communications (Takasaki *et al.*, 1976 and Yamamoto, 1980). Usually, for direct detection systems, (Alexander, 1987) the modulation formats used are: On-Off Keying (OOK), Pulse Position Modulation (PPM), and Manchester coded OOK. For coherent communications, the most commonly used are Frequency Shift Keying (FSK), Phase Shift Keying (PSK), and Amplitude Shift Keying (ASK). The choice of specific modulation format is dependent on the type of laser used, the target cost of the system, the type of channel present, and the performance goals of the system. Some signal sets convey information with a higher efficiency than others (Alexander, 1997). One way of specifying system efficiency is in terms of the number of photons required to transmit a bit of information at a specified probability of error, typically 10^{-9} . Increased power efficiency is usually accompanied by an increase in modulator and demodulator complexity.

JPL has chosen to use Pulse Position Modulation (PPM) for deep space optical communications. The PPM scheme was first proposed by Pierce in 1978 for the photon counting channel. In M-ary PPM, the data is divided into words that are $\log_2 M$ bits long. Each word corresponds to a specific transmitted symbol. Time is divided into M time slots, with an optical pulse sent in only one of the possible slots. It was demonstrated that the PPM scheme has unbounded transmission efficiency that can be increased infinitely

by merely increasing the number of pulse positions or time slots in the noiseless case. Since multiple bits of information can be transmitted using a single pulse, PPM can be more efficient than OOK in terms of the number of photons required to transmit an individual bit of information. PPM systems are also attractive when there is a substantial amount of optical background present. An extended description of this modulation scheme will be provided in future chapters of the thesis. PPM combined with forward error correcting codes can be used to obtain high transmission efficiencies at tolerable bandwidth expansion. The efficiency of the PPM format has led to its widespread consideration for the optical satellite communications channel. For this experiment, 32-PPM (5 bits per pulse) and 256-PPM (8 bits per pulse) systems have been proposed and their performance analyzed to be used with a heterodyne detection system. PPM systems convert average laser power to peak power directly proportional to M . Thus, the higher M systems will operate at higher K_s values (Gagliardi and Karp, 1995). 4-ary PPM systems have been demonstrated (Alexander, 1997) in some moderate data rate free-space systems (McGregor and Dion, 1991, and Sun *et al.*, 1990). The signaling format is very attractive as mentioned earlier because it becomes increasingly power efficient as the alphabet size increases. Large alphabets have been proposed for power efficient optical communications with space probes (Lesh, 1982). Unfortunately there are also some disadvantages to using PPM, including the need for accurate clock synchronization across all the time-slots (Sun and Davidson, 1990, Davidson and Sun, 1989), which becomes especially difficult with higher PPM orders. Another disadvantage of PPM is that semiconductor lasers are typically peak power limited, and when M increases the pulse

width decreases and the peak power increases. Therefore it is not feasible to achieve a high PPM order with semiconductor lasers (Katz, 1986), but in the experiments of this thesis, this is not a problem as already stated the selection of solid state lasers. Another problem that needs to be addressed is that as M increases, the receiver bandwidth increases as well, and it becomes harder to achieve low-noise performance with a wider bandwidth. This is why PPM is more commonly used in moderate data-rates systems (Davidson and Sun, 1991).

For the purpose of modulating the signal laser in the experiments, an external modulator configuration was selected. A PPM generator was used in conjunction with a CONOPTICS electro-optic (Pockle cell) modulator Model 350-210. It is a modulator of the transverse field type, that is, the electric field signal voltage is perpendicular to the optical propagation direction. The voltage swing required given operating wavelength to transit between the full off state to the full on state is called the Half Wave Voltage ($V_{1/2}$). The transverse field structure allows reduction of $V_{1/2}$ by manipulation of the crystal length to aperture ratio to a level achievable by available driver electronics. $V_{1/2}$ is roughly proportional to wavelength, and long wavelength devices usually require higher length to aperture ratios to accommodate existing driver output levels. The modulator is constructed using crystal Potassium Dideuterium Phosphate (KD*P). For the wavelength used, 1064 nm, the right model would belong to the 360 series, but it was not available at JPL. The only option was to use the only available model, the 350-210, which is actually built for 530 nm. With this model there is more insertion loss, multiple reflections are present, and the dynamic range is only 37%, as out of the 240 V that are needed for the

modulator driver, only 100 V are used when operated at 1064 nm. More details will be provided in the following chapters.

The output of a laser transmitter is often collimated by a transmitter optical antenna to reduce beam divergence. Beam divergence, by the principle of diffraction, is inversely proportional to the antenna aperture diameter. Newtonian or Cassegrainian beam collimators could be used.

2.2.2 Optical channel

A free-space optical channel is under study in this thesis. More specifically the transmitter will be located in deep space, and the optical beam will travel to Earth and be distorted by atmospheric turbulence. Effects of turbulence include intensity fluctuations (scintillation), beam wander, and beam spread. Classical turbulence is commonly explained following the Kolmogorov theory of turbulence, which is the set of hypothesis that a small-scale structure is statistically homogeneous, isotropic, and independent of the large-scale structure. The source of energy at large scales is either wind shear or convection. When the wind velocity is greater than the critical Reynolds number, large unstable air masses are created (Andrews, 2004).

2.2.3 Receiver architecture

This is the fundamental component of the optical communications link under research in this thesis. The performance of an optical deep space link depends on receiver sensitivity measured in photons per bit. The goal is to optimize the receiver design in order to obtain the highest possible sensitivity.

As it was mentioned in the previous chapter, an optical receiver could be implemented as a direct detection receiver or a coherent receiver. For the case of a direct detection system, the implementation is not complex; basically the signal collected is filtered in order to reduce the amount of background noise incident on the photodetector. The capacity of this type of receiver has been studied by several authors. Pierce (1978) proposed the Pulse Position Modulation format as a means to improve channel capacity. More recent work by Wyner (1988) demonstrated that in the presence of background noise, the capacity of a direct detection receiver may be expressed as follows:

$$C = (\log_2 e) \frac{\lambda_s}{M} \left[\left(1 + \frac{1}{\rho} \right) \right] \ln(1 + \rho) - \left(1 + \frac{M}{\rho} \right) \ln \left(1 + \frac{1}{\rho} \right) \quad (2.2.1)$$

where λ_s is the rate of arrival of the detected signal photon (photons/sec), $\rho = \frac{\lambda_s}{\lambda_b}$ is the peak signal to background power (Lesh, 2005) ratio and M is the peak to average power ratio for certain values of the average signal to background noise ratios. For the case of a high peak to average power ratio (Lesh, Katz, Tan, Zwillinger, 1981, and Katz, 1982), a transmission efficiency of one bit per photon is attainable.

For the case of a coherent system as previously explained, the signal coming from deep space is mixed with a strong local oscillator (LO) and the combined signal reaches the photodetector surface. This process assures linear amplification and downconverts (heterodyne) the optical signal into an electrical output with a frequency of a few MHz in the case to be presented in this thesis. If the local oscillator field is strong compared to the incoming signal field, it is possible to “amplify” the signal power above the noise

power of the electronics that constitute the front-end of the receiver and this allows shot noise limited performance to be achieved. Another advantage of this type of receiver is that it is much less sensitive to background noise since only the contribution from the background noise that is in the same spatial-temporal mode of the LO will be detected. Hence, coherent detection is a perfect candidate for operation with very high levels of background noise as would be the case if the Sun were in the FOV (field of view).

In terms of capacity, the sensitivity of a coherent receiver may be expressed as (C. Chen, 2005)

$$C = (\log_2 e) B \ln \left(1 + \frac{\lambda_s}{B} \right) \approx (\log_2 e) \lambda_s \quad (2.2.2)$$

where λ_s is the rate of detected signal photons, and the approximated final result assumed a large bandwidth B . As a result, from Eq. (2.2.2), the maximum attainable capacity of a heterodyne optical channel is ~ 1.44 bits/photon (Yamamoto and Haus, 1986). This type of receiver is more difficult to implement than a direct detection receiver since almost perfect spatial-mode matching has to be achieved. As an important goal of this thesis is to obtain the highest possible sensitivity, which is possible using a heterodyne receiver that will perform very close to the quantum limit; this is the type of system that will be investigated.

A heterodyne detection system was selected for deep space optical communications with noncoherent (or asynchronous) demodulation since this choice should improve receiver sensitivity by 3 dB compared to that of a direct detection system (Gagliardi, 1995). The choice of noncoherent demodulation is motivated to a high degree

by the physics of the problem. A difficulty that arises with PPM modulation is that no signal is present except in the signal slots. This makes it infeasible, in some but not all cases, to use a phase-locked loop (PLL) to track a signal carrier for two reasons. The loop can lose lock when the PPM pulse is over, and the duration of a PPM pulse may not give the loop adequate time to reacquire lock. Hence, the most common technique for maintaining phase synchronization in communications systems often will not apply. Furthermore, the wavelength of 1064 nm considered here is only slightly longer than 1 micron. When the optical path between the transmitter and the receiver goes through a change in length of one wavelength, there is a phase change of 2π radians. Since the distance from a remote spacecraft to Earth changes very quickly, the phase from one PPM pulse to the next is an i.i.d. uniform random variable over the range $(-\pi, \pi]$. This makes traditional phase tracking methods such as PLLs difficult to use in this case unless extremely accurate position and velocity predicts are used as well, and non-coherent detection of the received signal is the most realistic option in many cases.

Even low relative velocities of just 100 meters per second will result in a phase ramp that will go through a full cycle of 2π radians approximately 10^8 times per second. Assuming a data rate of 100 Mbps, this corresponds to a full phase cycle once every bit period. For the case of 256 PPM, the symbol rate would be 12.5 megasymbols/sec, yielding a total of 8 phase cycles for each PPM symbol. Due to uncertainties in spacecraft absolute position and velocity, the phase from one pulse to the next will be an i.i.d. uniform random process as described above, and errors in the predicted speed and

position would have to be very low to avoid this issue. If a signal were continuously present, as in BPSK or QPSK modulation, a tracking loop could be used to track the phase and enable fully coherent demodulation. The absence of a signal except in the signal slot, however, makes coherent phase detection a daunting engineering problem that is not taken up in this thesis. For these reasons, a receiver that uses noncoherent or asynchronous demodulation is the most realistic choice in this case.

At this point, it is useful to examine the meaning of the term “coherent.” The optical receiver that is the subject of this work is optically coherent. Instead of performing photon counting, it mixes the incoming optical signal with a locally generated local oscillator laser to produce an intermediate frequency (IF) communications signal that can be demodulated and processed using traditional methods. Hence, the term “optically coherent” will be used in reference to systems that mix incoming laser signals with a local oscillator laser and demodulate the resulting IF signal. This term differs from the term “coherent” in the communications context. In communications systems, the term “coherent” is used in reference to those systems that are able to track the phase of the incoming signal. Since this is not always a feasible task, a non-coherent envelope detection communications receiver was chosen.

The choice of a focal plane array based receiver instead of a single detector based receiver was driven by the presence of atmospheric turbulence. Even in clear weather there are random, time-varying patterns in the index of refraction of the air in the Earth’s atmosphere. When a laser signal from a distance spacecraft (assumed to be a plane wave due to the tremendous distances involved in deep space communications) passes through

the atmosphere, the differences in refractive index will distort the planar phase front of the incoming signal wave. This causes the wavefront that reaches the receiving telescope aperture to suffer from significant phase distortions that will cause the received signal pattern in the receiver focal plane to be spread out and be significantly distorted. A single small detection element designed to capture the Airy pattern resulting from a perfect plane wave reaching the aperture will be too small to catch most of the signal, resulting in severe signal losses. One could attempt to compensate for the spreading of the signal in the detector focal plane by using a single very large receiving element in the focal plane. However, the field in the focal plane is a complex field with both magnitude and phase. A single large detector element will effectively perform a simple addition of the focal plane field over its surface, and this could easily result in *destructive* signal cancellation, again resulting in severe signal losses. The key to capturing the scattered, phase and amplitude distorted signal field in the focal plane is to have an array of small elements. As signal components of varying amplitude and phase hit the array, the outputs of the receiving elements are coherently added together in order to reconstruct the original signal, avoiding both the self-cancellation that often occurs with large detection elements and the waste of useful signal energy in the focal plane that results from having a single small detection element incapable of capturing the entire signal field.

The selection of the size and geometry of a coherent optical focal plane array will be driven by a series of key factors:

1. The wavelength λ of the signal laser.
2. The diameter D of the telescope aperture.

3. The focal length f of the telescope.
4. The Fried parameter r_0 that characterizes turbulence. This parameter is described in greater detail below.

It is assumed that the incoming signal wavefront is of unit amplitude and also that it suffers from phase distortions induced by atmospheric turbulence following Kolmogorov's model (Andrews, 2004). This turbulence is characterized by a coherence length r_0 . It is generally assumed that the phase does not experience RMS fluctuations of more than one radian over a distance r_0 in the incoming wavefront. A low value of r_0 implies an atmospheric phase screen with great phase changes over a short scale. A high value of r_0 implies an atmospheric phase screen with high spatial correlation whose phase varies slowly as a function of distance.

It is well known that the approximate diameter of the Airy pattern in the focal plane of a telescope is $\frac{\lambda f}{D}$ (based on the Fourier transform relationship between the aperture plane and focal plane fields (Goodman, 2005) if the wavefront reaching the telescope aperture is an ideal plane wave. It is also known that the average size of the spread and distorted focal plane field in the presence of Kolmogorov turbulence is approximately $\frac{\lambda f}{r_0}$. Hence, the presence of atmospheric turbulence will increase the diameter of the signal field in the focal plane by a factor of approximately $\frac{D}{r_0}$. For example, a one-meter diameter telescope operating with $r_0 = 0.1$ m will see a tenfold

increase in its focal plane spot diameter. The need for a focal plane array to capture all of this energy for coherent recombining is now clear.

The need for *adaptive* coherent combining algorithms is a direct result of the fact that the random refractive index fluctuations in the atmosphere that spread and distort the focal plane field are rapidly time-varying fluctuations that must be tracked by an adaptive algorithm.

The nucleus of an optical communication receiver is the opto-electronic device that is used as the photodetector. An ideal photodetector would detect a 100% of all incident photons, would respond to the fastest changes on the incoming signal that were of interest (Alexander and Stephen, 1997), and would not introduce any additional noise besides the inherent quantum shot-noise from the received signal. The desired photodetector should be small, lightweight, cost effective, reliable, and it should not age with time and environmental conditions. The reality is that photodetectors have limited bandwidths with finite response times. The probability of detecting an individual photon is less than 100%, and unwanted noise is introduced in the photodetection process. Depending on the technology used, some photodetectors are more sensitive to aging or environmental conditions.

The photon-effect based photodetectors used in optical communications directly generate photocurrents from interactions between photons and atoms of the detector material. When a photodetector material is illuminated, a great number of atoms interact with the light, but the probability that a specific atom will absorb a photon and generate free carriers in order to form a photocurrent is small but dependent on the quality of the

design of the detector. For high-data rate communications, it is important to design detectors with fast response times and high sensitivities. The four kinds of detectors that could be used are: photomultipliers, photoconductors, photodiodes, and avalanche photodiodes. Characteristics of each of them (Alexander and Stephen, 1997) could be further explained, but it is a subject out of the scope of this thesis, and only a brief summary will be included.

A photomultiplier tube (PMT) is a form of vacuum tube that utilizes the photoelectric effect and the secondary emission of electrons to provide high current gains once a photon is initially detected. A photoconductor works using photon absorption in semiconductor materials. The p - n photodiode (PD) type uses a p - n junction to form a diode (called a photodiode). It relies on the absorption of photons in a semiconductor material. The photogenerated carriers are separated by an applied electric field, and the resulting photocurrent is proportional to the incident optical power. There is a variation of the p - n photodiode, which is the p - i - n photodiode that consists on a p - n photodiode with a layer of intrinsic or lightly doped (to assure that its characteristics are well controlled) semiconductor located between the p and n layers. Another type of photodiode is the avalanche photodiode or APD, capable of generating many electron-hole pairs in response to the absorption of a photon. APDs are good candidates for sensitive receivers because of their internal gain.

The SNR ratio is often the most important factor in choosing a detector (Kaufmann, 2005). The noise in a detector has components that are proportional to the signal and those that are independent of the signal. The components that are independent

of the signal consist of the thermal noise, or Johnson noise, and the dark current noise found in all types of detectors. The Johnson noise (i_j) is found only in solid-state detectors such as APDs and photodiodes (PDs), and can be described by

$$i_j = (4kTB / R_{sh})^{1/2} \quad (2.2.3)$$

where k is Boltzmann's constant, T is the absolute temperature, B is the frequency bandwidth, and R_{sh} is the shunt resistance.

Dark current noise, or dark-current shot noise (i_{sd}), may be defined as

$$i_{sd} = (2qI_d B)^{1/2} \quad (2.2.4)$$

where q is the electron charge and I_d is the dark current. The signal also has a noise component that, just like the dark current noise, is caused by statistical fluctuations. It is given by the corresponding formula

$$i_{ss} = (2qI_s B)^{1/2} \quad (2.2.5)$$

where I_s is the photocurrent created in the detector when the light is converted via the photoelectric effect into electrons. For a simple detector, the SNR could be expressed as

$$\begin{aligned} \text{SNR} &= \frac{I_s}{(i_j^2 + i_{sd}^2 + i_{ss}^2)^{1/2}} \\ &= \frac{I_s}{\left[(4kTB / R_{sh}) + (2qI_d B) + (2qI_s B) \right]^{1/2}} \end{aligned} \quad (2.2.6)$$

If the bandwidth is reduced, SNR could be improved. In the case of relatively high signal levels at the detector, SNR can be obtained basically by the contribution of the shot noise of the photocurrent; and then the SNR becomes proportional to $(I_s)^{1/2}$. This

regime is known as the shot-noise limit and represents ideal behavior for the detector. This limit when I_s is much greater than $0.26/R_{sh}$ and I_s is much greater than I_{sd} .

At high signal intensity, when the photocurrent is large enough to be in the shot-noise limit, the best choice of detector is a PD. The final choice of device will be driven by the wavelength and a cost tradeoff between optics that reduce the size of the light spot that falls on the detector and the cost per square millimeter for the active area of the detector. For the wavelength range from 200 to 1100 nm, the best choice is a silicon detector. For the near-IR (NIR) spectral region (1000 to 2600 nm), indium-gallium-arsenide (InGaAs) detectors provide a good solution. Between 2600 and 10,000 nm, a variety of detectors are available, including thermopiles and devices made of lead sulfide, lead selenide, indium arsenide, or mercury cadmium telluride. The photodetector that best suits the requirements of the receiver under study is the InGaAs p-i-n diode (with wavelength of 1064 nm) as it has a high efficiency on the order of 70-90%, a uniform spectral response, unity gain, and good stability. Furthermore, there are models that are quite fast.

Once the photodetector type has been selected, the front-end design has to be studied in order to decide on the best option. There are four types of front-end designs for a receiver (Alexander, 1997); resistor termination with a low-impedance voltage amplifier, transimpedance amplifier, high-impedance amplifier and noise-matched or resonant amplifier. The front-end of the receiver will respond to the optical signal generating a photocurrent with the *p-i-n* photodiode, then, this photocurrent will be converted to a voltage. The following step is to perform electrical signal processing in

order to extract the desired information from the recovered voltage. The dimensions of the transfer function associated with this front-end stage will be in volts per amp or ohms.

The configuration selected is the transimpedance amplifier (Alexander, 1997) as it provides a good compromise between the low-noise characteristics of the high-impedance front-end and the wideband nature of the low impedance voltage-amplifier. It is called a transimpedance amplifier because there is a shunt feedback around an inverting amplifier, which is a technique that is known to stabilize the amplifier's transimpedance (Holt, 1978). Figure 9 shows a transimpedance amplifier front end. Since the feedback utilizes the transimpedance and since the transfer function of an optical front end is inherently a transimpedance, this configuration is quite useful in optical communication receivers (Hullet and Muoi, 1976).

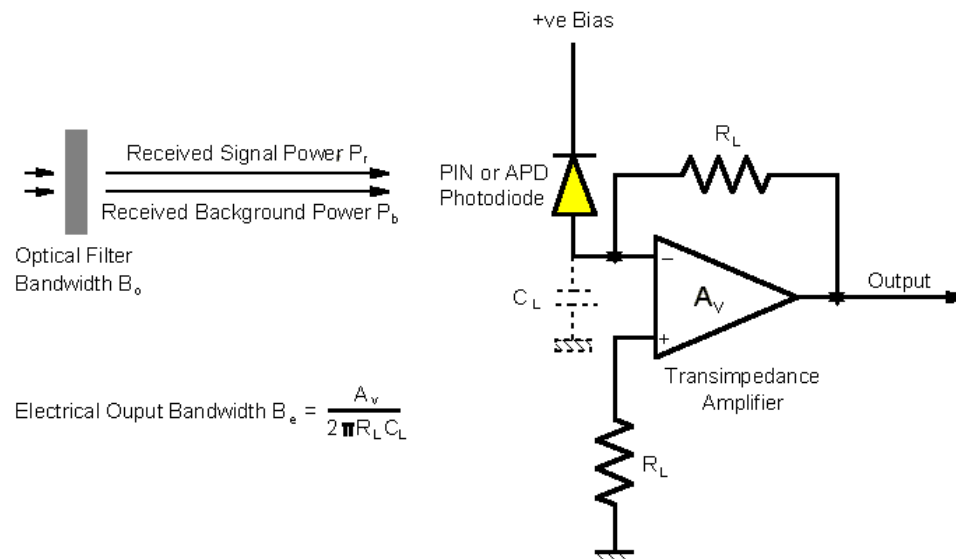


Figure 9. Generic transimpedance front-end receiver configuration.

As a conclusion on this subject, the transimpedance amplifier represents a compromise between the wideband resistor terminated low-impedance voltage amplifier (but comparatively noisy) approach, and the high-impedance type with low-noise, but equalization dependent (Alexander, 1997).

The photoreceiver selected was New Focus Model 1611-AC-FS (free space), a 1 GHz bandwidth photoreceiver based on an InGaAs PIN photodiode with a diameter of 0.1 mm. It includes an integrated lens (1.5 mm diameter ball lens) for quick optical coupling to the photodetector. It offers a well-balanced combination of gain, bandwidth and very low noise. It is AC-coupled with a low-frequency roll-off at 30 kHz. Figure 10 shows the functional block diagram of the 1611 photoreceiver (New Focus Manual),

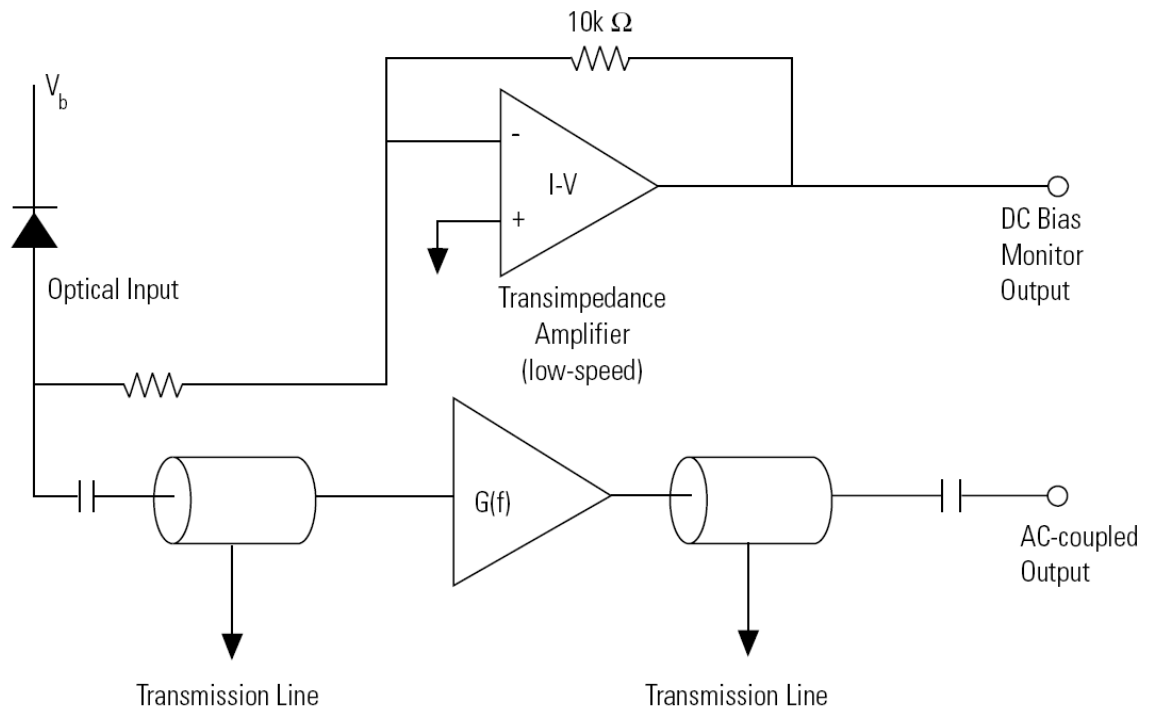


Figure 10. Functional diagram of the 1611 photoreceiver.

In order to do low-power measurements, a photodetector with high sensitivity or high conversion gain is needed. Conversion gain is wavelength dependent and is directly proportional to the responsivity. Therefore, in order to calculate the conversion gain at the operating wavelength, 1064 nm, the only thing that is needed is the variation of responsivity as a function of wavelength. Figure 11 shows the responsivity of the photodiode used in Model 1611. Responsivity (R) is the amount of photocurrent (I_{photo}) that results from an optical input of 1 W. This number can be used to calculate the photocurrent that will result from the experiment's input power (P_{in}) using the formula,

$$I_{photo} = RP_{in} \quad (2.2.7)$$

Responsivity is related to the quantum efficiency (the number of electrons released by an incident photon) by

$$QE = \eta_q = \frac{Rh\nu}{e} \quad (2.2.8)$$

where h is Planck's constant, ν is the frequency of the incident radiation, and e is the electron's charge.

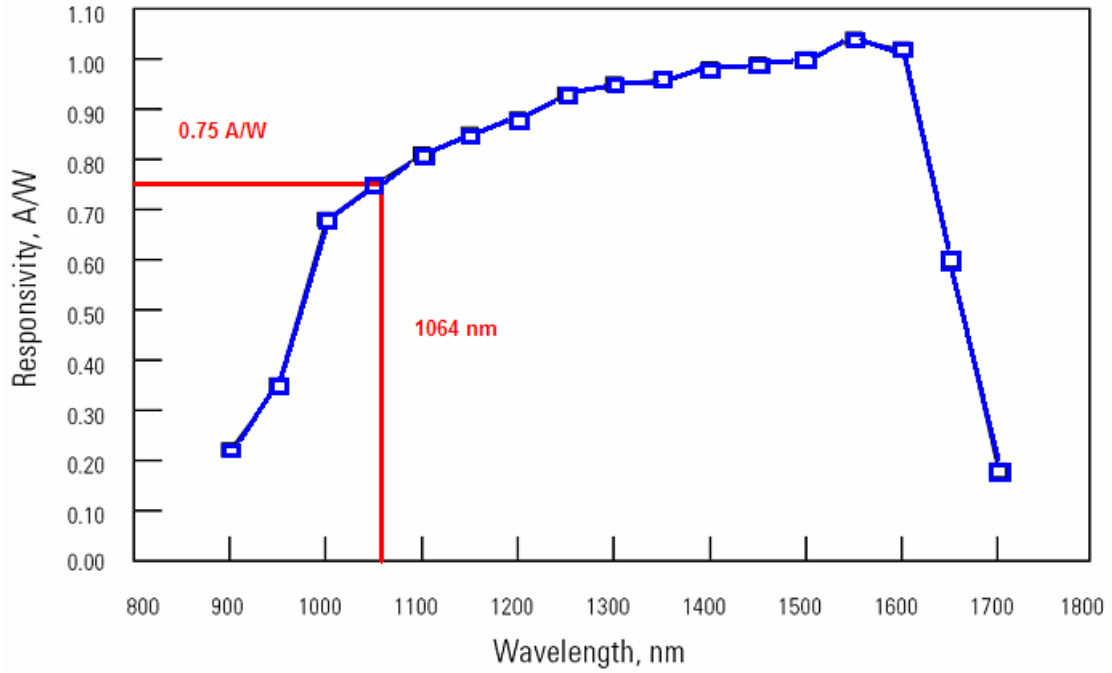


Figure 11. Responsivity of 1611 New Focus InGaAs photodetector.

Therefore, as shown in Figure 11, the responsivity for a wavelength of 1064 nm is 0.75 A/W. The quantum energy $h\nu = 1.16607$ eV, and using Eq. (2.2.8) the quantum efficiency computed is 87% for this type of detector.

Regarding implementation of a coherent receiver in general, a good option would be to use a balanced receiver configuration. There are two main advantages; first of all, since the balanced receiver generates the difference between the noise processes in the two photodetectors, this results in a cancellation of any correlated noise terms that appear in the two detectors. Intensity fluctuations that appear due to RIN in the local oscillator laser (LO) are correlated in the two photodetectors and are cancelled (Abbas *et al.*, 1985). The signal shot noises are uncorrelated in the two photodetectors and are not cancelled.

In order to achieve good wideband noise cancellation, it is critical to obtain a stable and accurate balance (Alexander and Stephen, 1997) between both photocurrents and also to have equal paths from the optical coupler to the each photodetector (Abbas *et al.*, 1985 and Alexander, 1987). In practice, it is easy to achieve shot noise limited performance using this configuration (Henry and Personick, 1990 and Abbas *et al.*, 1985). Another advantage of using the balanced receiver configuration is that it allows efficient use of LO and signal power as all the available power can be used instead of being wasted as in the case of the single detector receiver, where part of the LO as well as the signal laser power will be discarded.

In the experiments in this thesis, it was not possible to implement the balanced receiver configuration due to equipment and budget constraints. For single channel experiments, the 1611 New Focus photodetector was used. For the multichannel configuration, a focal plane array of 16 InGaAs detectors was designed at JPL and fabricated by Fermionics Opto-Technology. It is a 4 X 4 array of detectors that allows individual channels access for high speed signal processing. The individual detector elements have large active areas with minimal shadowing from bonded leads (approximately 250 μm diameter active area) as shown in Figure 12.

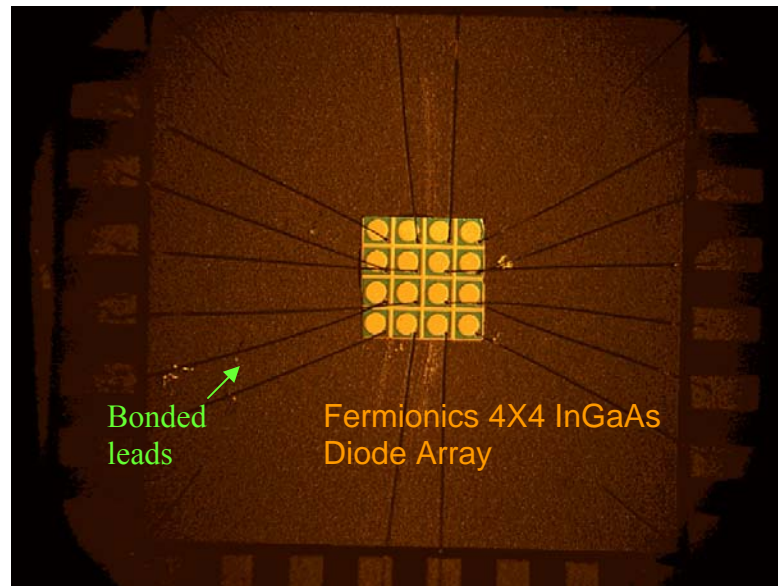


Figure 12. Detailed view of the 4 X 4 FPA of InGaAs detectors.

For the focal plane array, a set of 16 MITEQ AU-1447 amplifiers (200 MHz, 56 dB) were used. This corresponds to the previously mentioned configuration of resistor termination with a low-impedance voltage amplifier that unfortunately presents high thermal noise, but only one transimpedance amplifier was available. It was used for the single channel coherent receiver system that achieves nearly shot noise limited performance as shown later in the thesis (5 dB above thermal noise).

Chapter 3: Performance Analysis of the Coherent Optical Receiver Array

3.1 Introduction

Figure 13 shows the heterodyne receiver under study where the optical field is projected onto the photodetector surface by the receiver lens and front-end system. A local field is aligned by means of a mirror and combined with a beamsplitter with the received field in the photodetector.

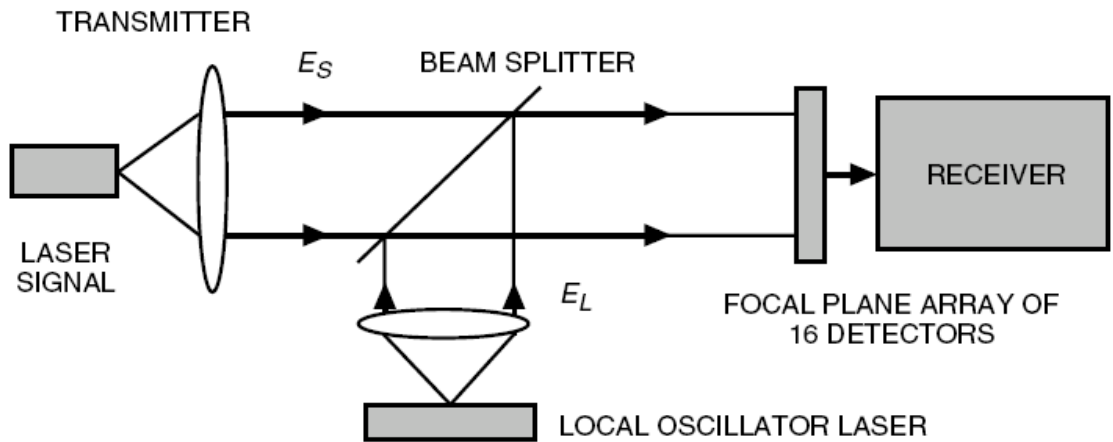


Figure 13. Diagram of the Coherent Optical Receiver Experiment (CORE).

The received field envelope of a heterodyne optical system may be described as (Gagliardi, 1995)

$$a_R(t) = s(t) + b(t) \quad (3.1.1)$$

where $s(t)$ and $b(t)$ are the complex envelopes of the signal and background noise respectively at frequency ω_s . The precise form of the signal envelope and its associated frequency spectrum will depend on the signal modulation. The received field mixes with a local oscillator field $a_L e^{j(\omega_L t + \theta_L)}$ at the beamsplitter. Assuming a photodetector of unity gain, the intensity per unit receiver area may be expressed as (Yariv, 1997)

$$I(t) \propto |a_{\mathbf{R}}(t)|^2 + |a_L|^2 + 2|a_L||a_{\mathbf{R}}(t)|\cos[(\omega_s - \omega_L)t + \theta_s(t) - \theta_L], \quad (3.1.2)$$

over the receiver area A . The intensity function is integrated over the receiver surface to yield the count intensity $c(t)$

$$c(t) = c_{\mathbf{R}}(t) + c_L(t) + c_{\mathbf{RL}}(t) \quad (3.1.3)$$

where

$$c_{\mathbf{R}}(t) = \alpha \int_A |a_{\mathbf{R}}(t)|^2 d\mathbf{r} = \alpha A |a_{\mathbf{R}}(t)|^2 \quad (3.1.4)$$

$$c_L(t) = \alpha \int_A |a_L|^2 d\mathbf{r} = \alpha A |a_L|^2 \quad (3.1.5)$$

$$\begin{aligned} c_{\mathbf{RL}}(t) &= \alpha \int_A 2|a_L||a_{\mathbf{R}}|\cos[(\omega_s - \omega_L)t + \theta_{\mathbf{R}}(t) - \theta_L] d\mathbf{r} \\ &= 2\alpha A |a_L||a_{\mathbf{R}}|\cos[(\omega_s - \omega_L)t + \theta_{\mathbf{R}}(t) - \theta_L] \\ &= 2\alpha A |a_L||a_s(t)|\cos[(\omega_s - \omega_L)t + \theta_s(t) - \theta_L] \\ &\quad + 2\alpha A |a_L||b(t)|\cos[(\omega_s - \omega_L)t - \theta_L] \end{aligned} \quad (3.1.6)$$

where $\alpha = \eta_q / h\nu$

In a typical heterodyne system, the power P_L of the local field will be much stronger than the received signal power ($a_L \gg a_s$). The strong local field condition assures a high count rate at the detector output even for very low power received signal fields. The current has an average value due entirely to the local field. The shot noise spectrum at the detector output is

$$N_{sh} = \alpha e^2 P_L, \quad (3.1.7)$$

where P_L is the average power of the local field,

$$P_L = a_L^2 A. \quad (3.1.8)$$

Because of the high count rate, the shot noise is a Gaussian process. Heterodyned detector outputs are almost always assumed to be Gaussian processes, with the signal term corresponding to the modulated carrier and all the other components considered as additive Gaussian noise (shot noise, background noise, and thermal noise), (Gagliardi, 1995).

Taking into account the heterodyned carrier energy in a symbol time T_s

$$E_s = 2(e\alpha)^2 P_s P_L T_s, \quad (3.1.9)$$

and the heterodyned noise level, including the two sided spectral components due to shot noise, background noise, and thermal noise, respectively may be expressed as (Gagliardi, 1995):

$$N_0 = e^2 [\alpha P_L + \alpha^2 P_L N_{ob}] + N_{th}, \quad (3.1.10)$$

As a result, the heterodyned decoding will be based on

$$\frac{E_s}{N_0} = \left[\frac{2\alpha P_s T_s}{1 + \alpha N_{ob} + (N_{th} / \alpha P_L e^2)} \right] \quad (3.1.11)$$

Thermal noise can be eliminated, and shot-noise limited performance can be achieved by means of a strong local field such that $\alpha P_L \gg N_{th} / e^2$. The local source power P_L produces an effective amplification of the received signal field as far as achieving shot-noise limited behavior. From Eq. (3.1.11), it can be observed that even if a strong local source is used, it is not possible to eliminate the term due to background noise. Introducing $K_s = \alpha P_s T_s$ as the detected average signal count per symbol, and assuming shot-noise limited operation due to a high local oscillator, Eq. (3.1.11) becomes

$$\frac{E_s}{N_0} = \left[\frac{2K_s}{1 + \alpha N_{ob}} \right] \quad (3.1.12)$$

In space systems with background noise such that $\alpha N_{ob} \ll 1$, Eq. (3.1.12) represents approximately a quantum-limited result expressed as

$$\frac{E_s}{N_0} = 2K_s \quad (3.1.13)$$

that can be considered a heterodyning quantum-limited bound.

3.2 Probability of bit error for coherent (heterodyne and homodyne) optical receiver for pulse position modulated (PPM) signals

When coherent detection is used, digital bits can be encoded directly on the phase or frequency of the laser carrier itself. The received modulated laser carrier can be translated to a lower RF frequency, where the digital modulation can be decoded using standard decoding techniques (Gagliardi and Karp, 1995).

In the heterodyne detection system examined, pulse position modulation (PPM) is used. PPM is a form of block encoding in which bits are transmitted in blocks instead of one at a time. Optical block encoding is achieved by converting each block of k bits into one of $M = 2^k$ optical fields of transmission. At the receiver end, decoding of each block is performed by determining which one of the M fields is received per block time. For the PPM case, a PPM frame contains M slots and an optical pulse is placed in one of those M slots. The data word is determined based on the position of the optical pulse in the frame, an example of which is shown in Figure 14.

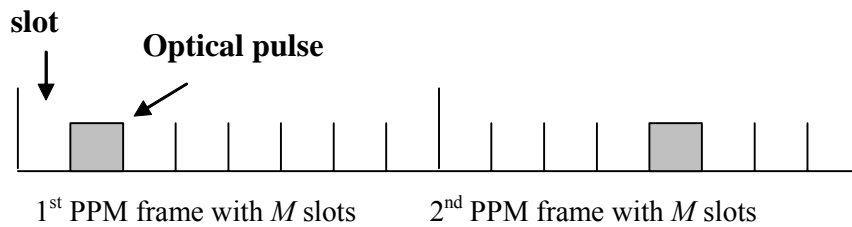


Figure 14. Pulse position modulation format.

The receiver decides on the basis of maximum likelihood symbol detection; it selects the slot with the greatest energy, and the symbol that contains a signal pulse in that slot location is declared to be the transmitted symbol. Under the assumptions of shot noise limited performance and negligible background noise levels, the performance of a coherent receiver will be examined in this section.

If A is the aperture of the receiver, and $a_L \gg a_s$, the resulting intensity counting rate process at the output of a unit gain photodetector is $\alpha\alpha e a_L^2 + 2\alpha e A a_s(t) a_L \cos[(\omega_s - \omega_L)t + \theta_s - \theta_L]$.

Under shot-noise limited conditions, and after filtering out the DC term, the detector can be modeled as:

$$r(t) = s(t) + n(t) \quad (3.2.1)$$

where $n(t)$ is a Gaussian noise process due to shot noise. The variance of the shot-noise is calculated by integrating for τ seconds (duration of the PPM pulse) resulting in $\sigma_s^2 = \alpha e^2 A a_L^2 \tau$. The value of the signal for homodyne detection is also obtained by integrating for τ seconds over the signal slot resulting in $2a_s a_L \alpha e A \tau$. For homodyne detection with $\omega_s = \omega_L$, and assuming perfect phase tracking, the resulting signal-to-noise ratio is

$$\begin{aligned} SNR &= \frac{[2a_s a_L \alpha e A \tau]^2}{\alpha e^2 A a_L^2 \tau} \\ &= 4\alpha A a_s^2 \tau \\ &= 4K_s \end{aligned} \quad (3.2.2)$$

where $K_s = \alpha A a_s^2 \tau$ is the average number of signal photons over the slot duration. For the case of heterodyne detection, the frequencies are not equal ($\omega_L \neq \omega_s$) and the signal becomes

$$s(t) = 2\alpha e A a_s a_L \cos[(\omega_s - \omega_L)t] \quad (3.2.3)$$

with rms value $\frac{2}{\sqrt{2}} a_s a_L \alpha e A \tau = \sqrt{2} a_s a_L \alpha e A \tau$. This results in the signal-to-noise ratio for heterodyne detection

$$\begin{aligned} SNR &= \frac{[\sqrt{2} a_s a_L \alpha e A \tau]^2}{\alpha e^2 A a_L^2 \tau} \\ &= 2\alpha A a_s^2 \tau \\ &= 2K_s \end{aligned} \quad (3.2.4)$$

The strong local field generates a high count rate at the detector output, which gives rise to Gaussian shot noise. Therefore, heterodyne detector outputs are assumed to be Gaussian processes with the signal term corresponding to the modulated carrier, and shot noise components considered as additive Gaussian noise with the spectral level given above. As a result, the photodetected field can be modeled as a Gaussian process, with mean $2a_s a_L \alpha e A \tau$ for homodyne detection and $\sqrt{2} a_s a_L \alpha e A \tau$ for heterodyne detection, and variance in both cases $\sigma_s^2 = \alpha e^2 A a_L^2 \tau$. The probability density can therefore be written as $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\eta)^2/2\sigma^2}$ where η is a mean value due to the signal energy.

The probability of correct PPM detection is the probability that one Gaussian random variable (Papoulis, 1991) with mean η exceeds $(M-1)$ other zero-mean Gaussian random variables. Since PPM signals are a type of orthogonal signals, for the homodyne detection case the probability of correct symbol detection $P(SC)$ can be expressed as (Viterbi, 1966):

$$P(SC) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi(\alpha e^2 a_L^2 A\tau)}} e^{-\frac{(x-2\alpha e a_L a_S A\tau)^2}{2(\alpha e^2 a_L^2 A\tau)}} dx \left[\int_{-\infty}^x \frac{1}{\sqrt{2\pi(\alpha e^2 a_L^2 A\tau)}} e^{-\frac{y^2}{2(\alpha e^2 a_L^2 A\tau)}} dy \right]^{M-1} \quad (3.2.5)$$

With the change of variables:

$$z = \frac{y}{\sqrt{\alpha e^2 a_L^2 A\tau}} \quad dz = \frac{dy}{\sqrt{\alpha e^2 a_L^2 A\tau}}, \text{ and noting that when } y = x, \quad z = \frac{x}{\sqrt{\alpha e^2 a_L^2 A\tau}}$$

the following simplified Eq. is obtained:

$$P(SC) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi(\alpha e^2 a_L^2 A\tau)}} e^{-\frac{(x-2\alpha e a_L a_S A\tau)^2}{2(\alpha e^2 a_L^2 A\tau)}} dx \left[\int_{-\infty}^{x/\sqrt{\alpha e^2 a_L^2 A\tau}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \right]^{M-1} \quad (3.2.6)$$

Then with another change of variables:

$$w = \frac{x}{\sqrt{\alpha e^2 a_L^2 A\tau}} \quad dw = \frac{dx}{\sqrt{\alpha e^2 a_L^2 A\tau}}, \text{ we get}$$

$$\begin{aligned} P(SC) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(w-\sqrt{4\alpha a_L^2 A\tau})^2}{2}} dw \left[\int_{-\infty}^w \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \right]^{M-1} \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(w-\sqrt{4K_S})^2}{2}} dw \left[\int_{-\infty}^w \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \right]^{M-1} \end{aligned} \quad (3.2.7)$$

or

$$P(SC) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(w-\sqrt{4K_S})^2}{2}} dw [1-Q(w)]^{M-1} \quad (3.2.8)$$

where

$$Q(w) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \quad (3.2.9)$$

Similarly, the probability of count symbol detection for the heterodyne case becomes

$$P(SC) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(w-\sqrt{2K_S})^2}{2}} dw [1-Q(w)]^{M-1} \quad (3.2.10)$$

If equal *a priori* transmission probabilities are assumed for each symbol, the probability of symbol error can be expressed as:

$$P_s(E) = 1-P(SC) \quad (3.2.11)$$

Finally, it is necessary to convert from symbol to bit errors (Simon *et al.*, 1995, Proakis, 2000). For the case treated in this derivation which is the study of orthogonal signals, the incorrectly decided symbol is equally likely to be any of the remaining $M-1$ symbols since each symbol is equidistant from all others. For equiprobable orthogonal signals, all symbol errors are equiprobable and occur with probability

$$\frac{P_s(E)}{M-1} = \frac{P_b(E)}{2^k-1} \quad (3.2.12)$$

There are $\binom{k}{n}$ ways in which n bits out of k may be in error. Therefore, the

average number of bit errors per k -bit symbol is:

$$\sum_{n=1}^k n \binom{k}{n} \frac{P_s(E)}{2^k - 1} = k \frac{2^{k-1}}{2^k - 1} P_s(E) \quad (3.2.13)$$

and the average bit error probability is obtained by dividing Eq. (3.2.13) by k , the number of bits per symbol. Hence,

$$\begin{aligned} P_b(E) &= \frac{2^{k-1}}{2^k - 1} P_s(E) \\ &= \frac{M/2}{M-1} P_s(E) \end{aligned} \quad (3.2.14)$$

Using Eq. (3.2.14) the probability of bit error for the case of homodyne detection is:

$$\begin{aligned} P_b(E) &= \frac{M/2}{M-1} \left[1 - \left\{ \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-(w-\sqrt{4K_s})^2/2} dw [1 - Q(w)]^{M-1} \right\} \right] \\ &= \frac{M/2}{M-1} \left[1 - \left\{ \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-(w-\sqrt{4K_s})^2/2} dw \left[\frac{1}{2} - \frac{1}{2} \operatorname{erf} \left(\frac{w}{\sqrt{2}} \right) \right]^{M-1} \right\} \right] \end{aligned} \quad (3.2.15)$$

For heterodyne detection the bit error probability can be expressed as:

$$P_b(E) = \frac{M/2}{M-1} \left[1 - \left\{ \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-(w-\sqrt{2K_s})^2/2} dw \left[\frac{1}{2} - \frac{1}{2} \operatorname{erf} \left(\frac{w}{\sqrt{2}} \right) \right]^{M-1} \right\} \right] \quad (3.2.16)$$

A simple bound often applied in block detection analysis is the union bound. The probability of a finite union of events is bounded above by the sum of the probabilities of

the constituent events. Since the binary test between any two decoding symbols is equivalent to an orthogonal coherent test, Eqs. (3.2.17) and (3.2.18) are obtained.

The union bound for the case of homodyne detection is:

$$P_b(E) \cong \left(\frac{M}{2}\right) \left[Q[\sqrt{2K_s}] \right] = \left(\frac{M}{2}\right) \left\{ \frac{1}{2} \operatorname{erfc}[\sqrt{K_s}] \right\} \quad (3.2.17)$$

Similarly, the union bound for the bit error probability for heterodyne detection becomes:

$$P_b(E) \cong \left(\frac{M}{2}\right) Q[\sqrt{K_s}] = \left(\frac{M}{2}\right) \left\{ \frac{1}{2} \operatorname{erfc} \left[\sqrt{\frac{K_s}{2}} \right] \right\} \quad (3.2.18)$$

The following figures show the exact bit error probabilities and union bound approximation for optical homodyne and heterodyne detection of PPM signals with $M = 2, 4, 8, 16$ slots. Note that as M increases the bit error probability is higher because we are plotting versus the average number of photons per pulse and not per bit. It can be observed that homodyne detection performs 3 dB better than heterodyne detection as expected.

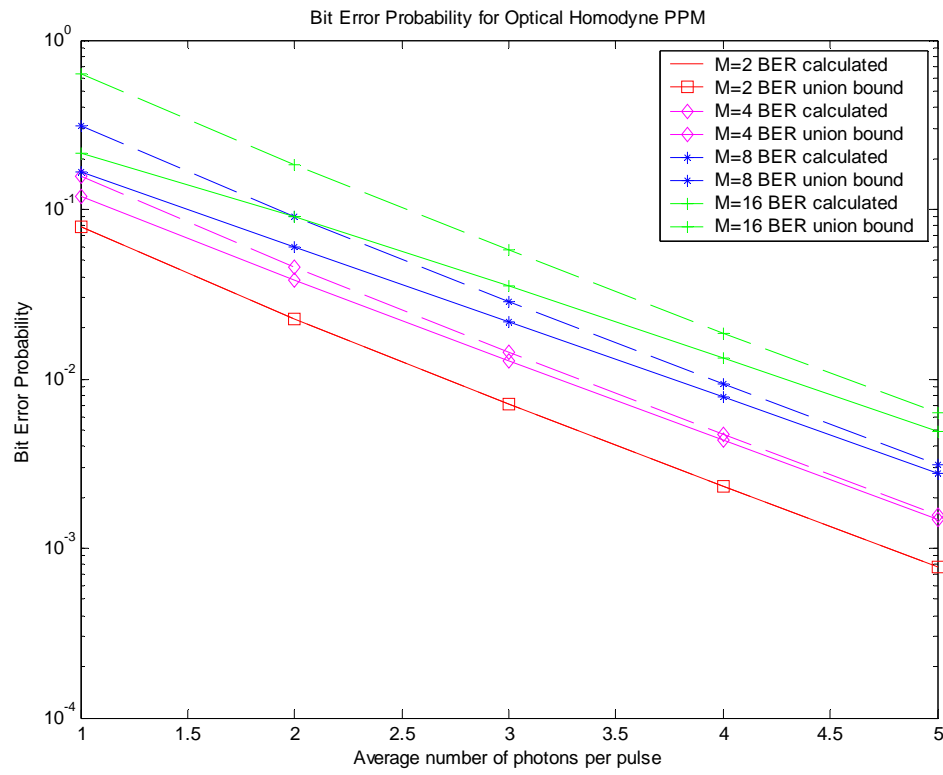


Figure 15. BER for homodyne detection.

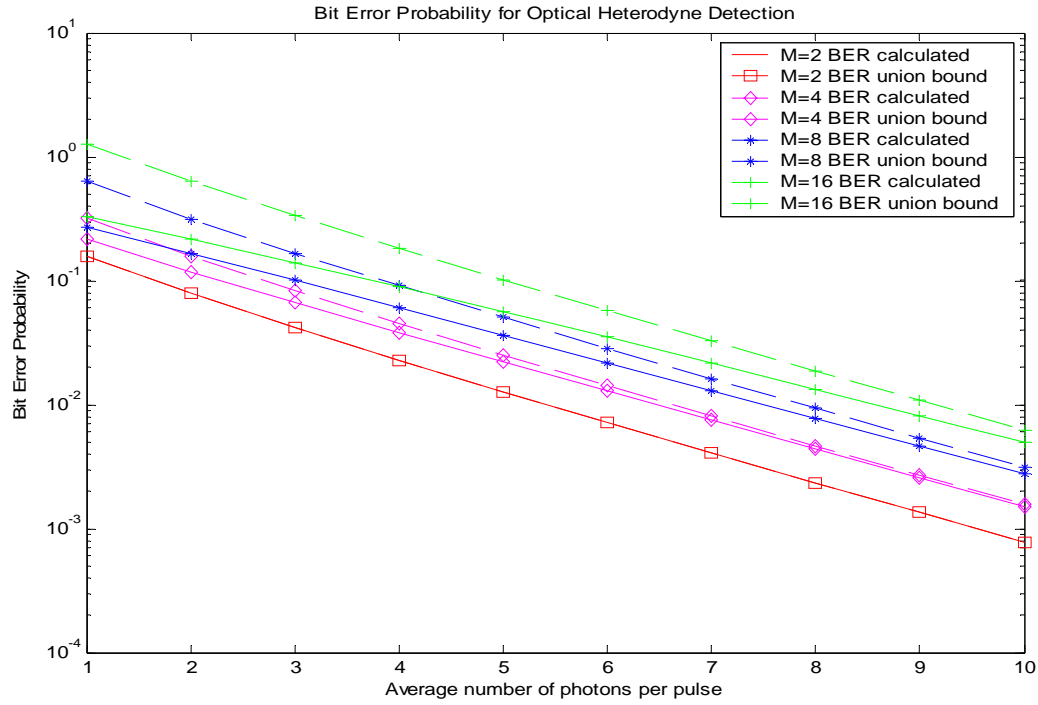


Figure 16. BER for heterodyne detection.

3.3 Probability of bit error for non-coherent (heterodyne) optical receiver with random phase channels for pulse position modulated (PPM) signals

3.3.1 Case I: Probability of bit error for the single channel

The receiver used here is non-coherent in the communications sense because the input phase is unknown and square-law envelope detection is used. As shown in (Gagliardi and Karp, 1995), the output process of a coherent optical heterodyning system at baseband is complex Gaussian. Since the phases of the complex Gaussian channel

outputs are not known, it is necessary to perform square law detection. The treatment here follows (Simon *et al.*, 1995). The real and imaginary parts of the complex signal are each Gaussian random variables. Summing the squares of two Gaussian random variables results in a chi-square random variable with two degrees of freedom. If both Gaussian random variables are zero mean, the resulting random variable at the output is a central chi-square random variable with two degrees of freedom.

For an M -ary pulse position modulation system, the probability of a correct symbol decision is the probability that none of the matched filter outputs in the noise slots exceeds the output in the signal slot. Let $f_{Y,sig}(y)$ be the probability distribution function (PDF) of the output within a signal slot, and let $F_{Y,noise}(y)$ be the cumulative distribution function (CDF) of the output in a noise slot. Three assumptions are made:

1. All M possible PPM symbols are equiprobable.
2. The PDF $f_{Y,sig}(y)$ always applies to any one of the M slots when the signal is present.
3. The CDF $F_{Y,noise}(y)$ always applies to each of the M slots when the signal is absent.

The probability of a correct decision is:

$$P(C) = \int_{-\infty}^{\infty} f_{Y,sig}(y) (F_{Y,noise}(y))^{M-1} dy \quad (3.3.1)$$

The symbol error probability is then given by:

$$\begin{aligned}
P(E) &= 1 - P(C) \\
&= 1 - \int_{-\infty}^{\infty} f_{Y, \text{sig}}(y) (F_{Y, \text{noise}}(y))^{M-1} dy
\end{aligned} \tag{3.3.2}$$

The probability of a bit error for M -ary PPM is related to the probability of symbol error by (Simon et al., 1995):

$$P_b(E) = \frac{M/2}{M-1} P(E) \tag{3.3.3}$$

The above Eqs. will be essential to the analysis of symbol error probabilities. Following the derivation on (Simon et al., 1995) for the case of equal energy signals the M -ary receiver selects the message m_k corresponding to maximum ξ_i^2 for $i = 0, 1, \dots, M-1$. The joint pdf $f_{\xi_0^2, \xi_1^2, \dots, \xi_{M-1}^2}(\xi_0^2, \xi_1^2, \dots, \xi_{M-1}^2)$ has to be obtained in order to compute the probability of symbol error. The received signal, which corresponds with the expression from Eq. (3.2.3) for a heterodyne system will be used in a generic form in this derivation in order to simplify the analysis. The received signal assuming message m_k is transmitted, is:

$$r(t) = \sqrt{2} s_k(t) \cos(\omega_c t + \theta) + n(t) \tag{3.3.4}$$

with $n(t)$ as the bandpass Gaussian noise process (Simon et al., 1995)

$$n(t) = \sqrt{2} \{n_c(t) \cos \omega_c t - n_s(t) \sin \omega_c t\} \tag{3.3.5}$$

where the baseband noise components are independent Gaussian processes with PSDs that may be expressed as

$$S_{nc}(f) = S_{ns}(f) = \begin{cases} \frac{N_0}{2} & |f| \leq B \\ 0 & \text{otherwise} \end{cases} \quad (3.3.6)$$

with

$$z_{ci} = \int_0^T r_c(t) s_i(t) dt = \int_0^T (s_k(t) \cos \theta + n_c(t) s_i(t)) dt \quad (3.3.7)$$

$$z_{si} = \int_0^T r_s(t) s_i(t) dt = \int_0^T (s_k(t) \cos \theta + n_s(t) s_i(t)) dt \quad (3.3.8)$$

The conditional means are computed for a fixed θ ,

$$E\{z_{ci} / \theta\} = \cos \theta \int_0^T (s_k(t) s_i(t)) dt = \begin{cases} E_s \cos \theta, & i = k \\ 0, & i \neq k \end{cases} \quad (3.3.9)$$

$$E\{z_{si} / \theta\} = \cos \theta \int_0^T (s_k(t) s_i(t)) dt = \begin{cases} E_s \sin \theta, & i = k \\ 0, & i \neq k \end{cases} \quad (3.3.10)$$

In order to compute the variances,

$$\sigma_{z_{ci}/\theta}^2 = E\left\{\left(\int_0^T n_c(t) s_i(t) dt\right)^2\right\} = \frac{N_0}{2} E_s \quad (3.3.11)$$

$$\sigma_{z_{si}/\theta}^2 = E\left\{\left(\int_0^T n_s(t) s_i(t) dt\right)^2\right\} = \frac{N_0}{2} E_s \quad (3.3.12)$$

Even if $n_c(t)$ and $n_s(t)$ are bandlimited processes; they could be considered to be nonbandlimited assuming that $BT \gg 1$, using the matched filter argument. As $n_c(t)$ and $n_s(t)$ are independent, z_{ci} and z_{si} are conditionally uncorrelated and therefore independent as they are conditionally Gaussian,

$$\begin{aligned}
f_{z_{ci}, z_{si}}(z, y / \theta) &= \frac{1}{\pi E_s N_0} \exp \left\{ -\frac{(z - E\{z_{ci} / \theta\})^2 + (y - E\{z_{si} / \theta\})^2}{E_s N_0} \right\} \\
&= \begin{cases} \frac{1}{\pi E_s N_0} \exp \left\{ -\frac{z^2 + y^2}{E_s N_0} \right\}, & i \neq k \\ \frac{1}{\pi E_s N_0} \exp \left\{ -\frac{z^2 + y^2 + E_s^2 - 2E_s z \cos \theta - 2yE_s \sin \theta}{E_s N_0} \right\}, & i = k \end{cases} \quad (3.3.13)
\end{aligned}$$

If the pairwise relation is considered among the various z_{ci} and z_{si} ,

$i = 0, 1, \dots, M-1$. Assuming that m_k is transmitted, then

$$\begin{aligned}
E\{z_{ci} z_{cj}\} &= E \left\{ \left[\int_0^T (s_k(t) \cos \theta + n_c(t) s_i(t) dt) \right] \left[\int_0^T (s_k(\tau) \cos \theta + n_c(\tau) s_i(\tau) d\tau) \right] \right\} \\
&= \left\{ \left(\int_0^T n_c(t) s_i(t) dt \right) \left(\int_0^T n_c(\tau) s_j(\tau) d\tau \right) \right\} \\
&= 0, \quad i \neq j, i \neq k, j \neq k \quad (3.3.14)
\end{aligned}$$

Also,

$$\begin{aligned}
E\{z_{ck} z_{cj}\} &= E \left\{ \left(E_s \cos \theta + \int_0^T n_c(t) s_k(t) dt \right) \left(\int_0^T n_c(\tau) s_j(\tau) d\tau \right) \right\} \\
&= 0, \quad j \neq k \quad (3.3.15)
\end{aligned}$$

Then,

$$\begin{aligned}
E\{z_{si} z_{sj}\} &= E\{z_{si}\} E\{z_{sj}\} = 0, \quad i \neq j \\
E\{z_{ci} z_{sj}\} &= E\{z_{ci}\} E\{z_{sj}\} = 0, \quad i = j \quad (3.3.16)
\end{aligned}$$

which is true for the cases of $i = k$ or $j = k$ but not both. As a result, the variables, z_{ci} and

z_{si} , $i = 0, 1, \dots, M-1$ are uncorrelated and therefore independent as they are

unconditionally Gaussian. Then if $\theta_i = \tan^{-1}(z_{si}/z_{ci})$, then Eq. (3.3.13) can be expressed as follows

$$f_{z_{ci}, z_{si}}(z, y/\theta) = \begin{cases} \frac{1}{\pi E_s N_0} \exp\left\{-\frac{z^2 + y^2}{E_s N_0}\right\}, & i \neq k \\ \frac{1}{\pi E_s N_0} \exp\left\{-\frac{z^2 + y^2 + E_s^2 - 2E_s \sqrt{z^2 + y^2} \cos(\theta - \theta_i)}{E_s N_0}\right\}, & i = k \end{cases} \quad (3.3.17)$$

Using the definition of the zero-order modified Bessel function of the first kind $I_0(x)$, then

$$f_{z_{ci}, z_{si}}(z, y) = \begin{cases} \frac{1}{\pi E_s N_0} \exp\left\{-\frac{z^2 + y^2}{E_s N_0}\right\}, & i \neq k \\ \frac{1}{\pi E_s N_0} \exp\left\{-\frac{z^2 + y^2 + E_s^2}{E_s N_0}\right\} I_0\left(\frac{2\sqrt{z^2 + y^2}}{N_0}\right), & i = k \end{cases} \quad (3.3.18)$$

Then, with the change of variables $\xi_i = \sqrt{z_{ci}^2 + z_{si}^2}$

$$f_{\xi_i}(\xi) = \begin{cases} \frac{2\xi}{E_s N_0} \exp\left\{-\frac{\xi^2}{E_s N_0}\right\}, & i \neq k \\ \frac{2\xi}{E_s N_0} \exp\left\{-\frac{\xi^2 + E_s^2}{E_s N_0}\right\} I_0\left(\frac{2\xi}{N_0}\right), & i = k \end{cases} \quad (3.3.19)$$

Normalizing ξ_i^2 and with

$$A_i = \frac{\xi_i^2}{E_s N_0} \quad (3.3.20)$$

then

$$f_{A_i}(a) = \frac{f_{\xi_i}(\sqrt{E_s N_0} a)}{2\sqrt{(a/E_s N_0)}} \quad (3.3.21)$$

Then,

$$f_{A_i}(a) = \begin{cases} \exp\left\{-a + \frac{E_s}{N_0}\right\} I_0\left(\sqrt{4\frac{E_s}{N_0}} a\right) u(a), & i = k \\ \exp\{-a\} u(a), & i \neq k \end{cases} \quad (3.3.22)$$

where $u(a)$ is the step function. The variables $A_i, i = 0, 1, \dots, M-1$ are independent because they are nonlinear functions of independent random variables. Then

$$f_{A_0, A_1, \dots, A_{M-1}}(a_0, a_1, \dots, a_{M-1}) = f_{A_0}(a_0) f_{A_1}(a_1) \dots f_{A_{M-1}}(a_{M-1}) \quad (3.3.23)$$

Using this expression for the joint pdf, it is possible to obtain the probability of symbol error $P_s(E)$. If message m_k is transmitted, and the output of the envelope detector matched to $s_k(t)$ is A_k , the probability of correct symbol detection conditioned on message m_k and A_k is

$$\begin{aligned} P_s(C / m_k, A_k) &= \text{Pr ob}\{A_0 \leq A_k, A_1 \leq A_k, \dots, A_{k-1} \leq A_k, A_{k+1} \leq A_k, \dots, A_{M-1} \leq A_k / m_k\} \\ &= \prod_{\substack{i=0 \\ i \neq k}}^{M-1} \int_{-\infty}^{A_k} f_{A_i}(a) da = (1 - e^{-A_k})^{M-1} \end{aligned} \quad (3.3.24)$$

$$\begin{aligned} P_s(C / m_k, A_k) &= \int_{-\infty}^{\infty} f_{A_k}(a_k) (1 - e^{-A_k})^{M-1} da_k \\ &= \exp\left\{-\frac{E_s}{N_0}\right\} \int_0^{\infty} e^{-a_k} I_0\left(\sqrt{4\frac{E_s}{N_0}} a_k\right) (1 - e^{-A_k})^{M-1} da_k \end{aligned} \quad (3.3.25)$$

Then, with the expansion

$$(1 - e^{-x})^{M-1} = \sum_{m=0}^{M-1} (-1)^m {}_{M-1}C_m e^{-mx} \quad (3.3.26)$$

where

$${}_{M-1}C_m = \frac{(M-1)!}{m!(M-1)!} \quad (3.3.27)$$

Then

$$P_s(C / m_k) = \exp\left\{-\frac{E_s}{N_0}\right\} \sum_{m=0}^{M-1} (-1)^m {}_{M-1}C_m e^{-mx} \times \int_0^\infty \exp\{-(m+1)a_k\} I_0\left(\sqrt{4\frac{E_s}{N_0}}\right) da_k \quad (3.3.28)$$

Taking into account that the integral

$$\int_0^\infty x \exp\{-a^2 x^2\} I_0(bx) dx = \frac{1}{2a^2} \exp\left\{\frac{b^2}{4a^2}\right\} \quad (3.3.29)$$

and with the change of variables $u = x^2$

$$\int_0^\infty \exp\{-a^2 u\} I_0(b\sqrt{u}) du = \frac{1}{a^2} \exp\left\{\frac{b^2}{4a^2}\right\} \quad (3.3.30)$$

The probability of correct symbol detection

$$P_s(C) = \sum_{m=0}^{M-1} (-1)^m {}_{M-1}C_m \frac{1}{m+1} \exp\left\{-\frac{m}{m+1} \left(\frac{E_s}{N_0}\right)\right\} \quad (3.3.31)$$

The probability of symbol error,

$$P_s(E) = 1 - P_s(C) \quad (3.3.32)$$

$$P_s(E) = \sum_{m=1}^{M-1} (-1)^{m+1} \frac{1}{m+1} {}_{M-1}C_m \exp\left\{-\frac{m}{m+1} \left(\frac{E_s}{N_0}\right)\right\} \quad (3.3.33)$$

As the expression for the probability of bit error rate is

$$P_b(E) = \frac{M/2}{M-1} P_s(E) \quad (3.3.34)$$

then

$$P_b(E) = \frac{M/2}{M-1} \sum_{m=1}^{M-1} (-1)^{m+1} \frac{1}{m+1} {}_{M-1}C_m \exp \left\{ -\frac{m}{m+1} \left(\frac{E_s}{N_0} \right) \right\} \quad (3.3.35)$$

with $E_s / N_0 = \text{SNR}$ of the channel under study.

The bit error rate for a non-coherent receiver in random phase channels could be calculated by substituting the signal-to-noise ratio obtained in Eq. (3.1.11).

3.3.2 Case II: Probability of bit error with array combining

For the case of signals under AWGN, the probability of bit error is given by the same expression from Eq. (3.3.35), but instead of using the SNR of one channel, the right term to use is the addition of the SNR of the individual channels to be combined (Vilnrotter *et al.*, 1992). Expressing the SNR directly as a ratio instead of in decibels and letting $\left(\frac{E_s}{N_0} \right)_i$ denote the SNR of the i^{th} channel, the total SNR is:

$$\left(\frac{E_s}{N_0} \right)_{\text{total}} = \sum_{i=1}^N \left(\frac{E_s}{N_0} \right)_i \quad (3.3.36)$$

for the case of N channels. Eq. (3.3.36) assumes that the noise is zero-mean Gaussian.

3.4 Probability of bit error for non-coherent (heterodyne) optical receiver with random phase channels for Pulse Position

Modulated (PPM) signals in the presence of leakage in the noise slots

The derivations from previous sections assume a system model with zero mean noise variances for signal as well as noise slots. Theoretically that is the case to be expected, but the reality is that the equipment used for this thesis had some limitations and constraints, necessitating changes in the mathematical model. The PPM modulator is not ideal, and during the dead time it was “leaking” signal energy. This unwanted signal energy, referred to as “leakage,” was mixed with the local oscillator laser, generating a heterodyned beatnote in the noise slots. This resulted in a non-zero signal mean in the noise slots, which in this section will be explained in more detail. Different channels presented different leakage characteristics that have been taken into account in the following signal model based on a non-central chi square distribution for the signal as well as the noise slots.

Previous sections analyzed the case of zero mean Gaussian random variables resulting in a central chi-square random variable with two degrees of freedom in the noise slots. If either or both of the Gaussian random variables has non-zero mean, then the result will be a non-central chi square random variable with two degrees of freedom. The central chi-square random variable is simply a special case of the non-central chi-square random variable (Simon *et al.*, 1995).

First, consider the matched filter output in the signal slot. Since the signal consists of two Gaussian random variables with non-zero mean in the real and imaginary

parts, the resulting pdf is a non-central chi-square pdf with two degrees of freedom. Let m_1 and m_2 be the means of the real and imaginary parts of the complex Gaussian, and let them both have variance σ^2 . Define the sum of the squares of the means:

$$s^2 = m_1^2 + m_2^2 \quad (3.4.1)$$

Then the pdf is given by:

$$f_{Y, sig}(y) = \frac{1}{2\sigma^2} \exp\left[-\frac{y+s^2}{2\sigma^2}\right] I_0\left(\frac{s\sqrt{y}}{\sigma^2}\right) \quad (3.4.2)$$

where $I_0(x)$ is the zeroth order modified Bessel function of the first kind.

If a leakage term is present in the noise slots, then the sum of squared means for those slots, s_L^2 , will be non-zero. The resulting cdf will then be:

$$\begin{aligned} F_{Y, noise}(y) &= 1 - Q_1\left(\frac{s_L}{\sigma}, \frac{\sqrt{y}}{\sigma}\right) \\ &= 1 - \int_{\frac{\sqrt{y}}{\sigma}}^{\infty} x \exp\left(-\frac{x^2 + \left(\frac{s_L}{\sigma}\right)^2}{2}\right) I_0\left(\frac{s_L}{\sigma} x\right) dx \end{aligned} \quad (3.4.3)$$

where $Q_1(a, b)$ is the Marcum-Q function.

Plugging Eqs. (3.4.2) and (3.4.3) into Eq. (3.3.1) yields:

$$P(E) = 1 - \frac{1}{2\sigma^2} \int_0^{\infty} \exp\left[-\frac{y+s^2}{2\sigma^2}\right] I_0\left(\frac{s\sqrt{y}}{\sigma^2}\right) \left(1 - Q_1\left(\frac{s_L}{\sigma}, \frac{\sqrt{y}}{\sigma}\right)\right)^{M-1} dy \quad (3.4.4)$$

The integral in Eq. (3.4.4) presents the following difficulties:

1. It does not have a known closed form expression.
2. It is difficult to evaluate numerically for large M (i.e., $M = 256$).

One approach is to upper bound and lower bound the Marcum-Q function, and by using numerical bounds judiciously it is possible to obtain upper and lower bounds for the probability of symbol error given in Eq. (3.4.4). The bounds on the Marcum-Q function are given by (Simon, 2002):

$$\exp\left[-\frac{(b+a)^2}{2}\right] \leq Q_1(a, b) \leq \exp\left[-\frac{(b-a)^2}{2}\right] \quad (3.4.5)$$

The bounds in Eq. (3.4.5) play an important role in the analysis of the bit error probability performance of this system. Another possibility involves direct approximation of the integral in Eq. (3.4.4). The first stage is to expand the probability of correct decision as shown below:

$$\begin{aligned} P(\text{Correct Decision}) &= \int_0^\infty P(\text{Correct Decision} | y) f_{Y, sig}(y) dy \\ &= \int_0^\infty (F_{Y, noise}(y))^{M-1} f_{Y, sig}(y) dy \\ &= \int_0^\infty (F_{Y, noise}(y))^{M-1} \left[\frac{1}{2\sigma^2} \exp\left(-\frac{y + s_{sig}^2}{2\sigma^2}\right) I_0\left(\frac{s}{\sigma^2} \sqrt{y}\right) \right] dy \\ &= \int_0^\infty \left(1 - Q_1\left(\frac{s_{leak}}{\sigma}, \frac{\sqrt{y}}{\sigma}\right)\right)^{M-1} \left[\frac{1}{2\sigma^2} \exp\left(-\frac{y + s_{sig}^2}{2\sigma^2}\right) I_0\left(\frac{s}{\sigma^2} \sqrt{y}\right) \right] dy_{sig} \end{aligned}$$

Next, a binomial expansion is followed by an approximation step:

$$\begin{aligned}
\left(1 - Q_1\left(\frac{s_{leak}}{\sigma}, \frac{\sqrt{y}}{\sigma}\right)\right)^{M-1} &= \sum_{m=0}^{M-1} \binom{M-1}{m} \left(-Q_1\left(\frac{s_{leak}}{\sigma}, \frac{\sqrt{y}}{\sigma}\right)\right)^m \\
&\approx 1 - (M-1)Q_1\left(\frac{s_{leak}}{\sigma}, \frac{\sqrt{y}}{\sigma}\right)
\end{aligned} \tag{3.4.6}$$

The approximation from Eq. (3.4.6) is usable if the following condition is met:

$$Q_1\left(\frac{s_{leak}}{\sigma}, \frac{\sqrt{y_{sig}}}{\sigma}\right) \ll \frac{1}{M-1} \tag{3.4.7}$$

The approximation of Eq. (3.4.6) is used to simplify the integral:

$$\begin{aligned}
&\int_0^\infty \left(1 - Q_1\left(\frac{s_{leak}}{\sigma}, \frac{\sqrt{y}}{\sigma}\right)\right)^{M-1} \left[\frac{1}{2\sigma^2} \exp\left(-\frac{y+s_{sig}^2}{2\sigma^2}\right) I_0\left(\frac{s}{\sigma^2} \sqrt{y}\right) \right] dy \\
&\approx \int_0^\infty \left[\frac{1}{2\sigma^2} \exp\left(-\frac{y+s_{sig}^2}{2\sigma^2}\right) I_0\left(\frac{s}{\sigma^2} \sqrt{y}\right) \right] dy - \\
&\quad (M-1) \int_0^\infty Q_1\left(\frac{s_{leak}}{\sigma}, \frac{\sqrt{y}}{\sigma}\right) \left[\frac{1}{2\sigma^2} \exp\left(-\frac{y+s_{sig}^2}{2\sigma^2}\right) I_0\left(\frac{s}{\sigma^2} \sqrt{y}\right) \right] dy \\
&= 1 - (M-1) \int_0^\infty Q_1\left(\frac{s_{leak}}{\sigma}, \frac{\sqrt{y}}{\sigma}\right) \left[\frac{1}{2\sigma^2} \exp\left(-\frac{y+s_{sig}^2}{2\sigma^2}\right) I_0\left(\frac{s}{\sigma^2} \sqrt{y}\right) \right] dy
\end{aligned} \tag{3.4.8}$$

The approximation used in Eq. (3.4.8) is very useful in that the integral on the last line has a closed form expression given by (Simon, 2002)

$$\begin{aligned}
&\int_0^\infty Q_1\left(\frac{s_{leak}}{\sigma}, \frac{\sqrt{y}}{\sigma}\right) \left[\frac{1}{2\sigma^2} \exp\left(-\frac{y+s^2}{2\sigma^2}\right) I_0\left(\frac{s}{\sigma^2} \sqrt{y}\right) \right] dy \\
&= Q_1\left(\frac{s_{leak}}{\sigma\sqrt{2}}, \frac{s}{\sigma\sqrt{2}}\right) - \frac{1}{2} \exp\left(-\frac{s^2+s_{leak}^2}{4\sigma^2}\right) I_0\left(\frac{s s_{leak}}{2\sigma^2}\right)
\end{aligned} \tag{3.4.9}$$

The theoretical BER error rate can therefore be approximated using Eqs. (3.4.9) and (3.4.8) to obtain:

$$\text{SER}(s, s_{leak}, \sigma) \approx (M-1) \left[Q_1 \left(\frac{s_{leak}}{\sigma\sqrt{2}}, \frac{s}{\sigma\sqrt{2}} \right) - \frac{1}{2} \exp \left(-\frac{s^2 + s_{leak}^2}{4\sigma^2} \right) I_0 \left(\frac{ss_{leak}}{2\sigma^2} \right) \right] \quad (3.4.10)$$

and

$$\begin{aligned} \text{BER}(s, s_{leak}, \sigma) &= \frac{M/2}{M-1} \text{SER}(s, s_{leak}, \sigma) \\ &\approx \frac{M}{2} \left[Q_1 \left(\frac{s_{leak}}{\sigma\sqrt{2}}, \frac{s}{\sigma\sqrt{2}} \right) - \frac{1}{2} \exp \left(-\frac{s^2 + s_{leak}^2}{4\sigma^2} \right) I_0 \left(\frac{ss_{leak}}{2\sigma^2} \right) \right] \end{aligned} \quad (3.4.11)$$

Below a series of expressions specific to the case without leakage are given. If there is no leakage, then $s_L = 0$ and the CDF $F_{Y,noise}(y)$ is then given by:

$$F_{Y,noise}(y) = 1 - Q_1 \left(0, \frac{\sqrt{y}}{\sigma} \right) \quad (3.4.12)$$

This makes the upper and lower bounds in (3.4.5) equal to each other and therefore tight.

It is now possible to write:

$$F_{Y,noise}(y) = 1 - \exp \left(-\frac{y}{2\sigma^2} \right) \quad (3.4.13)$$

The resulting expression for symbol error probability is:

$$1 - \frac{1}{2\sigma^2} \int_0^\infty \exp \left[-\frac{y + s^2}{2\sigma^2} \right] I_0 \left(\frac{s\sqrt{y}}{\sigma^2} \right) \left(1 - \exp \left(-\frac{y}{2\sigma^2} \right) \right)^{M-1} dy \quad (3.4.14)$$

which is equivalent to the expression derived by Simon *et al.* (1995) for M -ary orthogonal symbol error probabilities for non-coherent (from a communications perspective) receivers. Since:

$$\frac{E_s}{N_0} = \frac{s^2}{2\sigma^2} \quad (3.4.15)$$

is the SNR, it is possible to perform a change of variables in order to obtain symbol error probability as a function of SNR. The final answer in the no leakage case is (Simon et al., 1995):

$$P(E) = \sum_{m=1}^{M-1} \left((-1)^{m+1} \frac{1}{m+1} \binom{M-1}{m} \exp \left(-\frac{m}{m+1} \left(\frac{E_s}{N_0} \right) \right) \right) \quad (3.4.16)$$

3.4.1 Expressions specific to the case with leakage

Since even numerical evaluation of the integral for the symbol error probability in (3.4.4) is difficult in general, a different strategy was taken. An upper bound on $Q_1(a, b)$ yields a upper bound on the probability of making a symbol error in Eq. (3.4.17), and similarly a lower bound on $Q_1(a, b)$ will yield a lower bound on the symbol error probability. The upper bound expression is given by:

$$P_{upper}(E) = 1 - \frac{1}{2\sigma^2} \int_0^\infty \exp \left[-\frac{y+s^2}{2\sigma^2} \right] I_0 \left(\frac{s\sqrt{y}}{\sigma^2} \right) \left(1 - \exp \left(-\frac{(\sqrt{y}-s_L)^2}{2\sigma^2} \right) \right)^{M-1} dy \quad (3.4.17)$$

Similarly, the lower bound expression is:

$$P_{lower}(E) = 1 - \frac{1}{2\sigma^2} \int_0^\infty \exp \left[-\frac{y+s^2}{2\sigma^2} \right] I_0 \left(\frac{s\sqrt{y}}{\sigma^2} \right) \left(1 - \exp \left(-\frac{(\sqrt{y}+s_L)^2}{2\sigma^2} \right) \right)^{M-1} dy \quad (3.4.18)$$

The bounds on SER given by Eqs. (3.4.17) and (3.4.18) were evaluated via numerical integration using measured values of the following three parameters:

1. s^2 : The sum of the squared means of the matched filter outputs in a signal slot.
2. s_L^2 : The sum of the squared means of the matched filter outputs in a noise slot.
3. σ^2 : The variance of the matched filter real and imaginary parts. Since the real and imaginary parts each have variance σ^2 , the total variance of the complex Gaussian matched filter output is $2\sigma^2$.

The three parameters above must be estimated from the experimental data since the leakage is unknown *a priori*.

3.4.2 Obtaining the SNR, signal, noise, and leakage parameters

From Section 3.2, which follows the treatment in (Simon et al., 1995), it is known that given the time-varying input phase $\theta(t)$ of the incoming signal, the means of the Gaussian matched filter outputs are:

$$\begin{aligned} m_1(t) &= E_s \cos \theta(t) \\ m_2(t) &= E_s \sin \theta(t) \end{aligned} \tag{3.4.19}$$

The expression for s^2 is:

$$s^2 = m_1(t)^2 + m_2(t)^2 = E_s^2 \cos^2 \theta(t) + E_s^2 \sin^2 \theta(t) = E_s^2 \tag{3.4.20}$$

It is also known that the variances of the real and imaginary parts are given by:

$$\sigma^2 = \frac{E_s N_0}{2} \tag{3.4.21}$$

The SNR can be expressed as:

$$\begin{aligned}
\text{SNR} &= \frac{E_s}{N_0} \\
&= \frac{1}{2} \frac{E_s^2}{\left(\frac{E_s N_0}{2}\right)} \\
&= \frac{1}{2} \frac{m_1(t)^2 + m_2(t)^2}{\sigma^2} \\
&= \frac{s^2}{2\sigma^2}
\end{aligned} \tag{3.4.22}$$

The main difficulty in evaluating s^2 , s_L^2 , and σ^2 lies in the fact that the phase θ is really a slowly varying function of time $\theta(t)$. Since $\theta(t)$ is unknown, there is a need for a way to estimate the three key parameters s^2 , s_L^2 , and σ^2 . Given that the non-central chi-squared outputs are available, one can take advantage of two Eqs. that describe the mean and the variance of the non-central chi-square outputs, assuming two degrees of freedom (Simon et al., 1995). Let y be the non-central chi-square random variable. Then:

$$E[y] = 2\sigma^2 + s^2 \tag{3.4.23}$$

$$\text{Var}[y] = 4\sigma^4 + 4\sigma^2 s^2 \tag{3.4.24}$$

One computes the sample mean and the sample variance of the non-central chi-square outputs from the square-law device over all of the *signal* slot outputs. The results are plugged into Eqs. (3.4.23) and (3.4.24), and the Eqs. are solved to obtain both s^2 and σ^2 as shown below:

First, s^2 is expressed by solving Eq. (3.4.23) to obtain:

$$s^2 = E[y] - 2\sigma^2 \quad (3.4.25)$$

Next, plug Eq. (3.4.25) into Eq. (3.4.24) to obtain:

$$4\sigma^4 + 4\sigma^2(E[y] - 2\sigma^2) = \text{Var}[y] \quad (3.4.26)$$

Eq. (3.4.26) is then simplified to obtain a quadratic Eq. in σ^2 :

$$4(\sigma^2)^2 - 4\sigma^2 E[y] + \text{Var}[y] = 0 \quad (3.4.27)$$

where the expression $(\sigma^2)^2$ was used to make the quadratic form more apparent. The quadratic formula yields:

$$\sigma^2 = \frac{E[y] \pm \sqrt{(E[y])^2 - \text{Var}[y]}}{2} \quad (3.4.28)$$

Although Eq. (3.4.28) yields two solutions due to the \pm ambiguity, this ambiguity is readily resolved. Substituting Eq. (3.4.28) into Eq. (3.4.25) yields:

$$\begin{aligned} s^2 &= E[y] - 2 \frac{E[y] \pm \sqrt{(E[y])^2 - \text{Var}[y]}}{2} \\ &= E[y] - E[y] \mp \sqrt{(E[y])^2 - \text{Var}[y]} \\ &= \mp \sqrt{(E[y])^2 - \text{Var}[y]} \end{aligned} \quad (3.4.29)$$

Since s^2 has to be positive, one must chose the positive sign in the last line of Eq. (3.4.29). This implies choosing the negative sign in Eq. (3.4.28). The results are:

$$\sigma^2 = \frac{E[y] - \sqrt{(E[y])^2 - \text{Var}[y]}}{2} \quad (3.4.30)$$

and

$$s^2 = \sqrt{(E[y])^2 - \text{Var}[y]} \quad (3.4.31)$$

Although Eqs. (3.4.30) and (3.4.31) are mathematically true, the receiver does not have access to the true values of $E[y]$ and $\text{Var}[y]$. Instead, the receiver must compute time averages of these variables from the sample outputs of the matched filter for the signal slots. The time averages are themselves random variables that have the values of $E[y]$ and $\text{Var}[y]$ as their means. Since decision feedback must be used in a realistic receiver to determine which slot is most likely to be the signal slot, there will be some additional noise due to incorrect symbol decisions as well. This poses a significant issue

because the term $\sqrt{\left(\langle y \rangle_{\text{signal slots}}\right)^2 - \left\langle \left(y - \langle y \rangle_{\text{signal slots}}\right)^2 \right\rangle_{\text{signal slots}}}$ used to approximate the

term $\sqrt{(E[y])^2 - \text{Var}[y]}$ may be close to zero for very small values of s^2 . This is rarely a problem with the signal slots, but it is a significant problem with the noise slots if the leakage s_L^2 is very low because s_L^2 is computed just as s^2 is except that the averaging is performed over the noise slots and not over the signal slots. If the time average is less than zero, then the result of taking the radical would be an imaginary number. However, if the time average is less than zero, this usually signifies that the term s^2 (averaging over signal slots) or s_L^2 (averaging over noise slots) must be zero or very close to zero. In that case, one simplifies the problem by setting either s^2 or s_L^2 to zero as appropriate. One then solves Eq. (3.4.25) for σ^2 in a straightforward manner. The variance is the same for the noise slots as well as the signal slots. Once these temporal estimates are obtained, it

is possible to numerically compute the integrals in Eqs. (3.4.17) and (3.4.18) in order to bound the BER curves.

Summary

The methods used to obtain the SNR; the signal, noise and leakage statistics, and the upper and lower bounds on BER performance have been presented in this section. All of these operations are carried out in the software receiver. As an example, the following figures show coherently detected PPM beatnotes under atmospheric turbulence conditions (OFF time of the electro-optic modulator) where the leakage may be clearly observed in different channels during the noise slots.

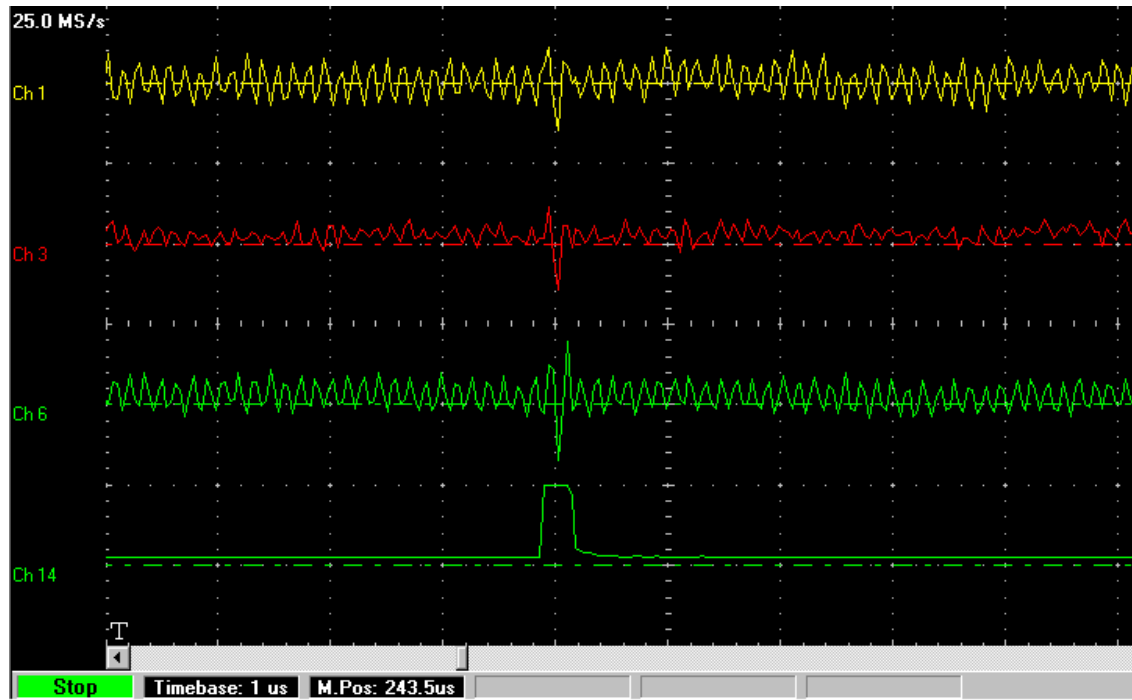


Figure 17. Coherently detected PPM beatnotes in the presence of atmospheric turbulence channels show leakage signal in noise slots example 1.

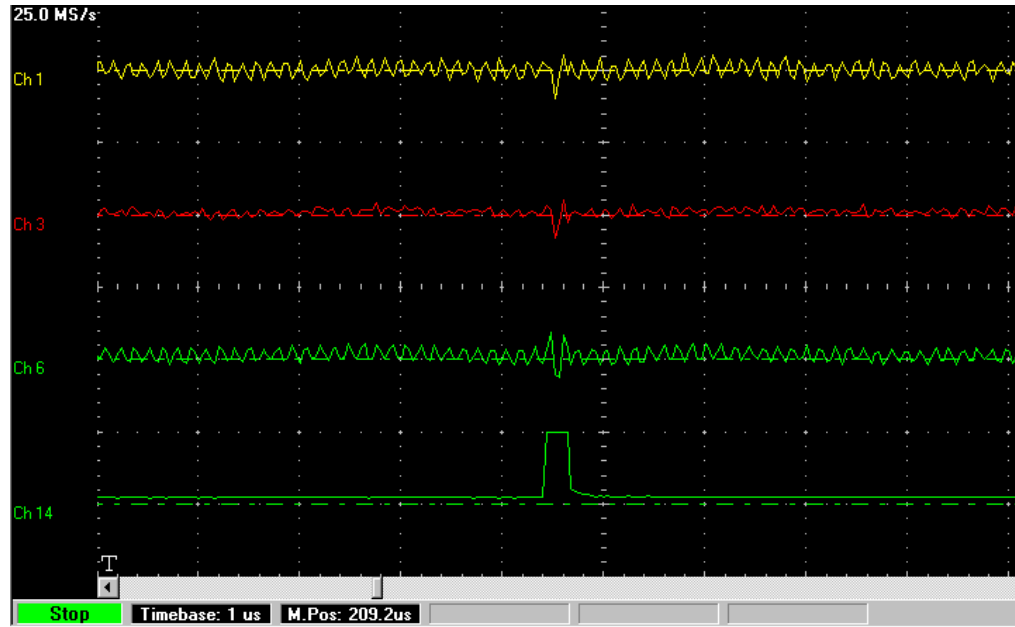


Figure 18. Coherently detected PPM beatnotes in the presence of atmospheric turbulence channels show leakage signal in noise slots example 2.

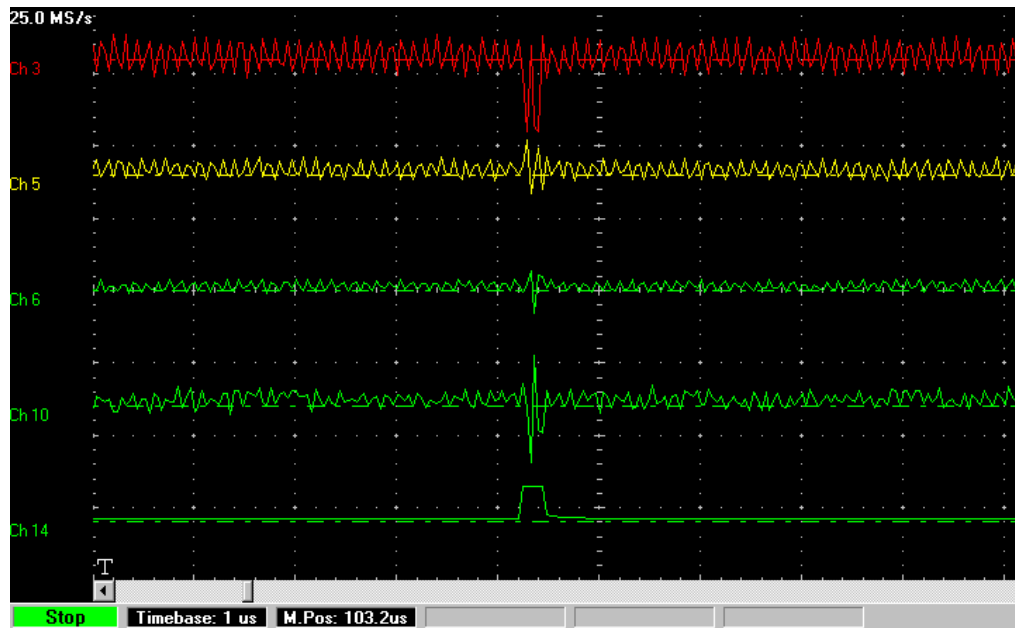


Figure 19. Coherently detected PPM beatnotes in the presence of atmospheric turbulence channels show leakage signal in noise slots example 3.

Chapter 4: Results of the Coherent Optical Receiver Experiment (CORE)

4.1 Description of the experimental setup

The experimental setup consists of two Nd:YAG (YAG denotes yttrium aluminum garnet) lasers operating at 1064 nm, whose outputs are aligned and combined on the surface of a 4 X 4 Fermionics InGaAs detector array. One of the lasers serves as a local oscillator, while the other simulates the received signal. The two lasers are operated at slightly different wavelengths, yielding a relatively stable difference-frequency tone of approximately 6 MHz in the detected signal. In the presence of spatial distortions simulating atmospheric turbulence conditions, the difference-frequency tone is generally observable in several array elements simultaneously, but usually with different phases. If the detector element outputs were simply summed, the addition of out-of-phase tones could result in significant cancellation, yielding a weak signal tone at the output. Non-coherent addition of signal components from different elements of the detector array is analogous to detection with a single large detector: this is the prime reason why a single large detector is not effective for coherent detection of signal fields under turbulent conditions. However, if small areas of the detector surface over which the signal field is essentially coherent are processed separately, then the outputs can be phase aligned prior to addition, recovering the lost signal power.

Figure 20 is a photograph of the optical setup at the Jet Propulsion Laboratory where the experiments described in this thesis have taken place. This signal laser beam is focused into the photodetector via the receiver lens.

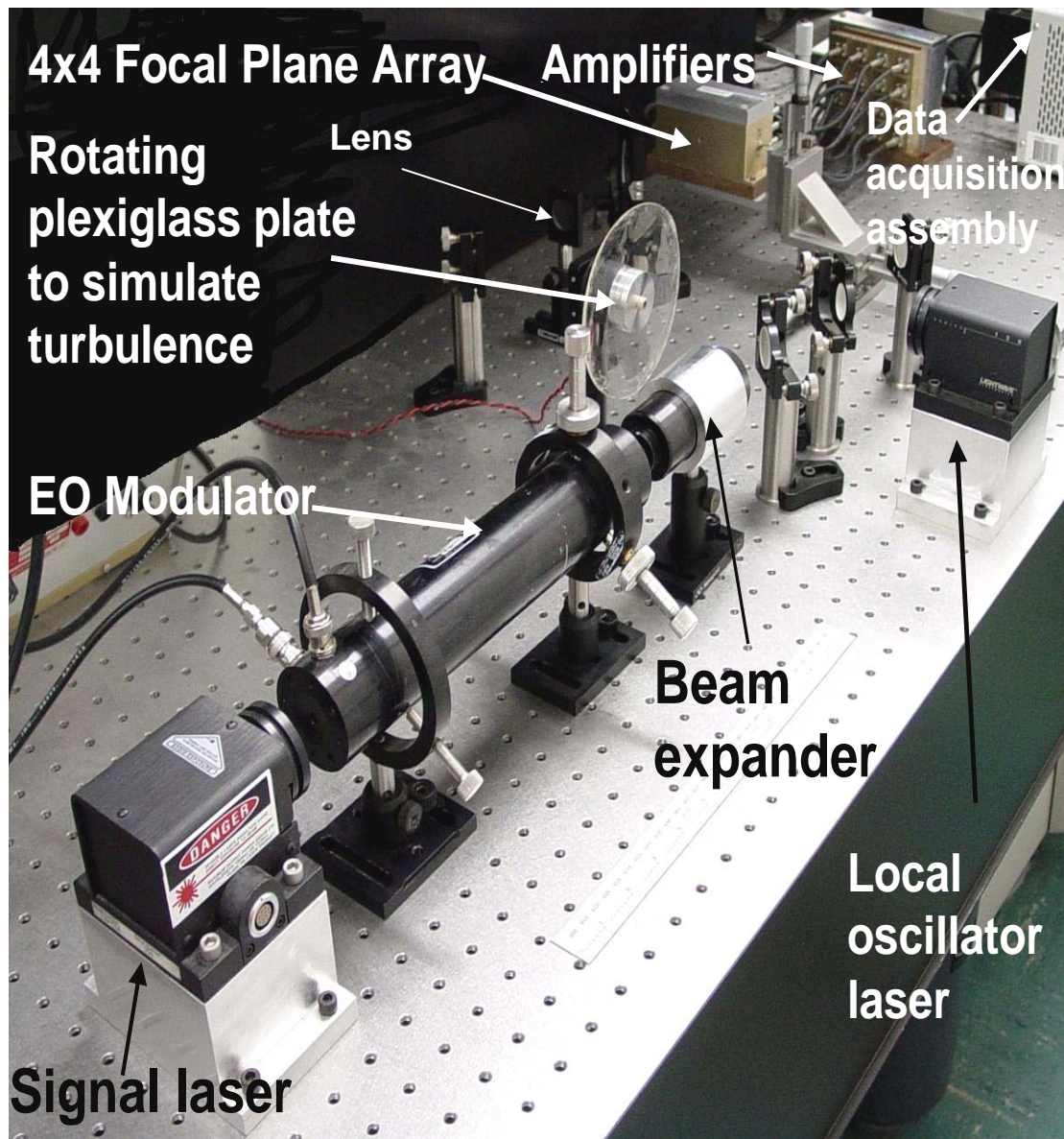


Figure 20. Coherent combining experiment at the Jet Propulsion Laboratory, NASA.

In the current coherent combining experiment, each of the 16 outputs of the detector array is amplified, and input to a 16-channel data-acquisition assembly (using GaGe data-acquisition cards). The analog signals are digitized to 8 bits at a sampling rate of 25 megasamples per second (MSPS). The data-acquisition system is capable of synchronously recording up to 1 megabyte of data per channel (or one million 8-bit samples). However for the initial tests, only 104,128 samples were taken per channel in order to simplify the data transfer from the data-acquisition computer to the signal-processing computer. At a sampling rate of 25 MSPS, this sample stream represents 4.16512 ms of elapsed time. For the first results, channels that contained significant signal were identified, and at a certain time synchronously collected 104,128 samples were collected from each channel (in a realistic communications scenario, the combining algorithm would automatically select the “signal” pixels for processing). The modulation format for the transmitted laser signal is PPM using an external Electro-Optic Modulator (Pockle cell). A snapshot of an individual laser pulse that contains the coherently detected PPM beatnote is shown in Figure 21. At the GaGe scope we can see the PPM modulated beatnotes (Figure 26) at a rate of 97.65 kHz. The PPM frame period is exactly 10.24 μ s. The slot width is approximately 320 ns, resulting in a 32 PPM system.

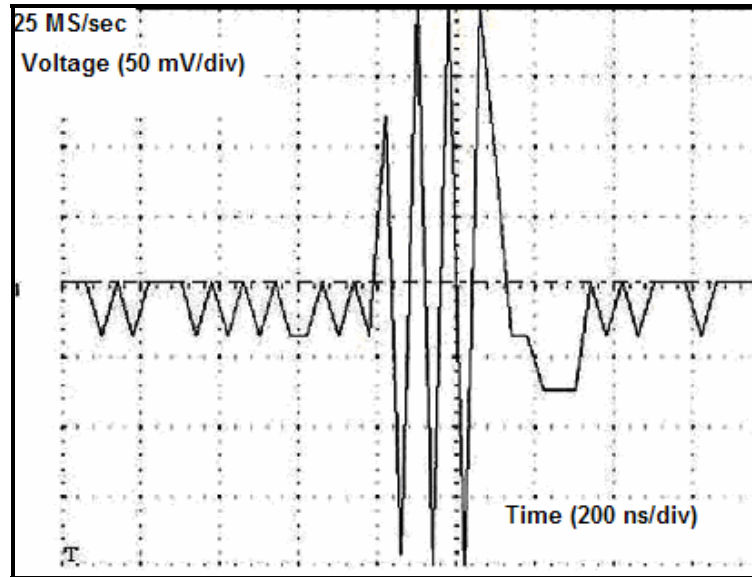


Figure 21. Snapshot of an individual PPM pulse beatnote.

The resulting sample stream acquired with the GageScope is digitally downconverted to complex baseband. The resulting downconverted complex samples served as input to a least-mean-square (LMS) algorithm, which was used to estimate the complex weights required to reconstruct the signal. The complex-weighted samples from each channel were then combined, in order to maximize the combined signal-to-noise ratio (SNR).

A rotating predistorted plexiglass plate was incorporated into the experimental setup to simulate atmospheric turbulence. Intensity distributions of the signal beam obtained with a Spiricon camera at the input to the focal-plane array under ideal conditions are shown in Figure 22.

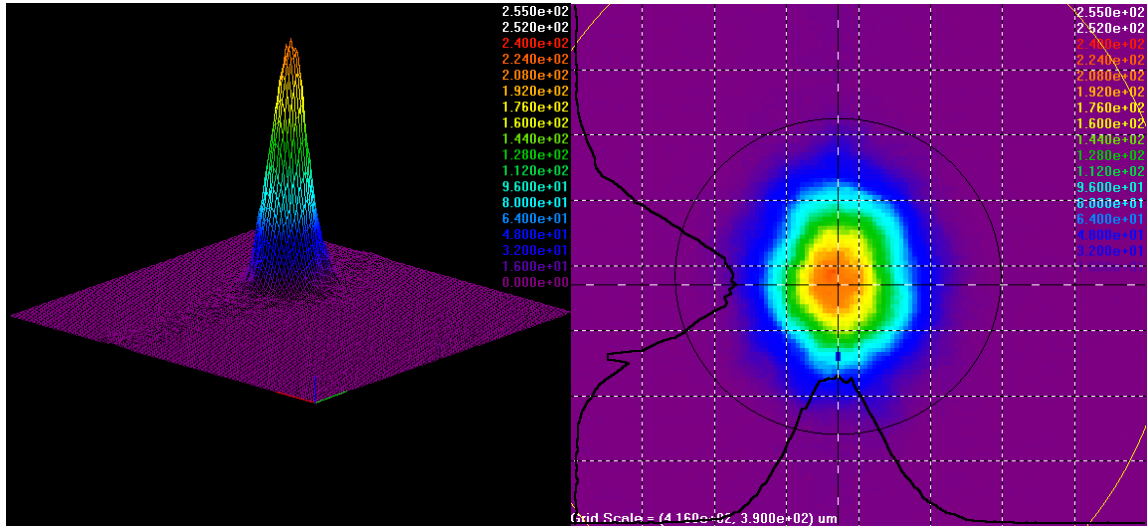


Figure 22. Intensity distribution of the signal beam under ideal conditions.

While when the plexiglass plate is introduced in the signal path, as observed in Figure 23, the beam is broken up into several “hot spots” that can be captured by different detectors of the focal plane array

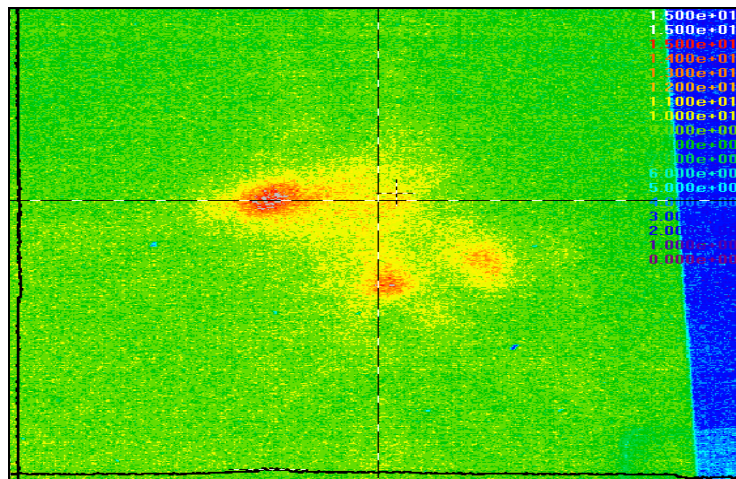


Figure 23. Beam profile under spatial distortions resembling atmospheric turbulence generated with a plexiglass plate in the laboratory.

The signal and local beams were characterized using the Spiricon camera; Figure 24 shows plots of the image taken at the detector surface for the signal and local oscillator beams respectively under ideal conditions, while Figure 25 shows perfect alignment of both beams with a 3D beam profile representation.

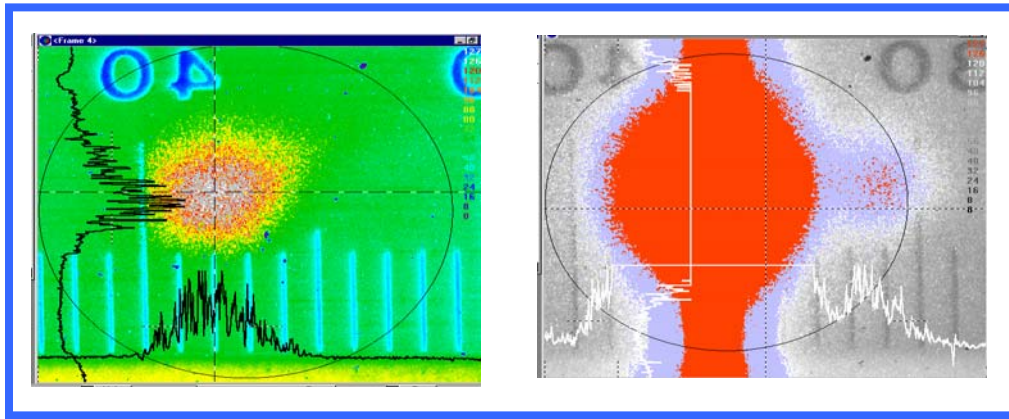


Figure 24. Beam profile of the signal laser (left) and local oscillator (right) beams on the detector surface.

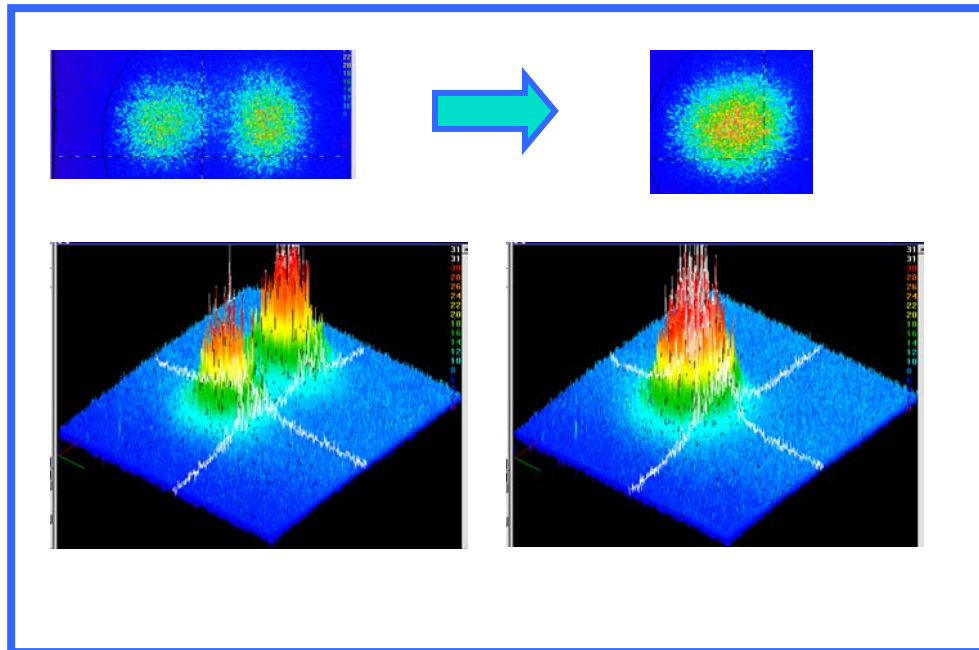


Figure 25. Perfect alignment of signal and local oscillator beams at the InGaAs detector surface.

Figure 26 illustrates coherently detected PPM beatnotes in four different channels under ideal conditions and Figure 27 shows the four channels in the presence of atmospheric attenuation.

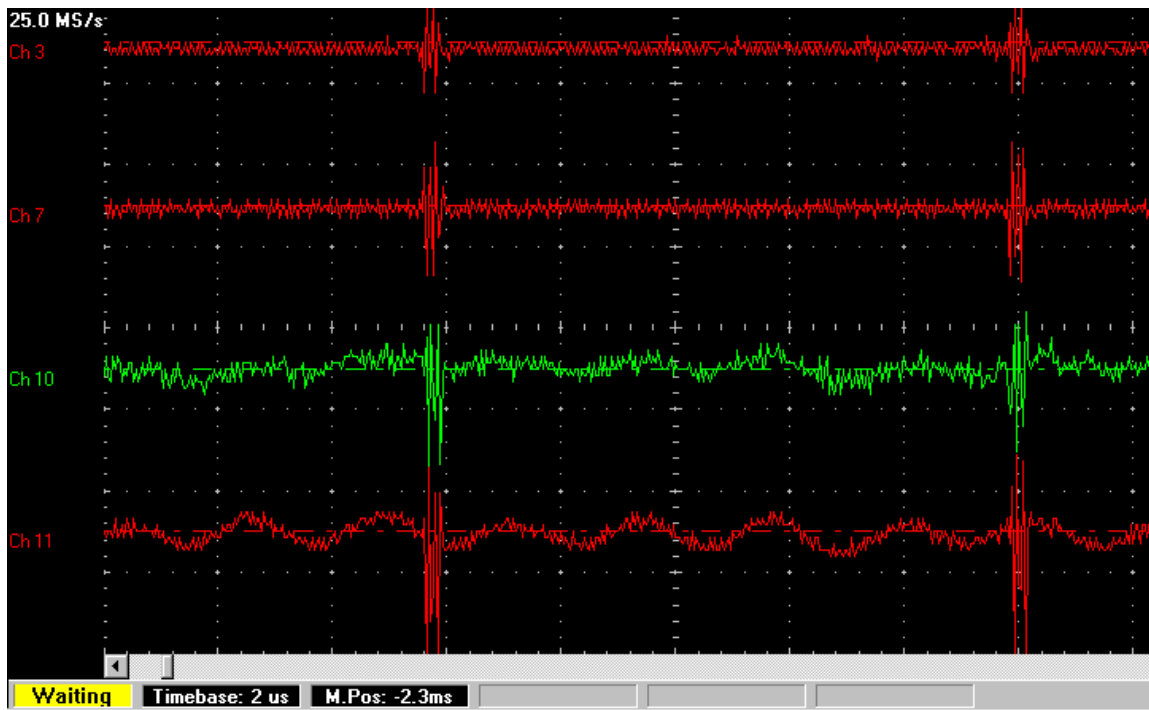


Figure 26. Sampled sequences of 4 channels containing PPM-modulated 6 MHz beatnotes under ideal conditions.

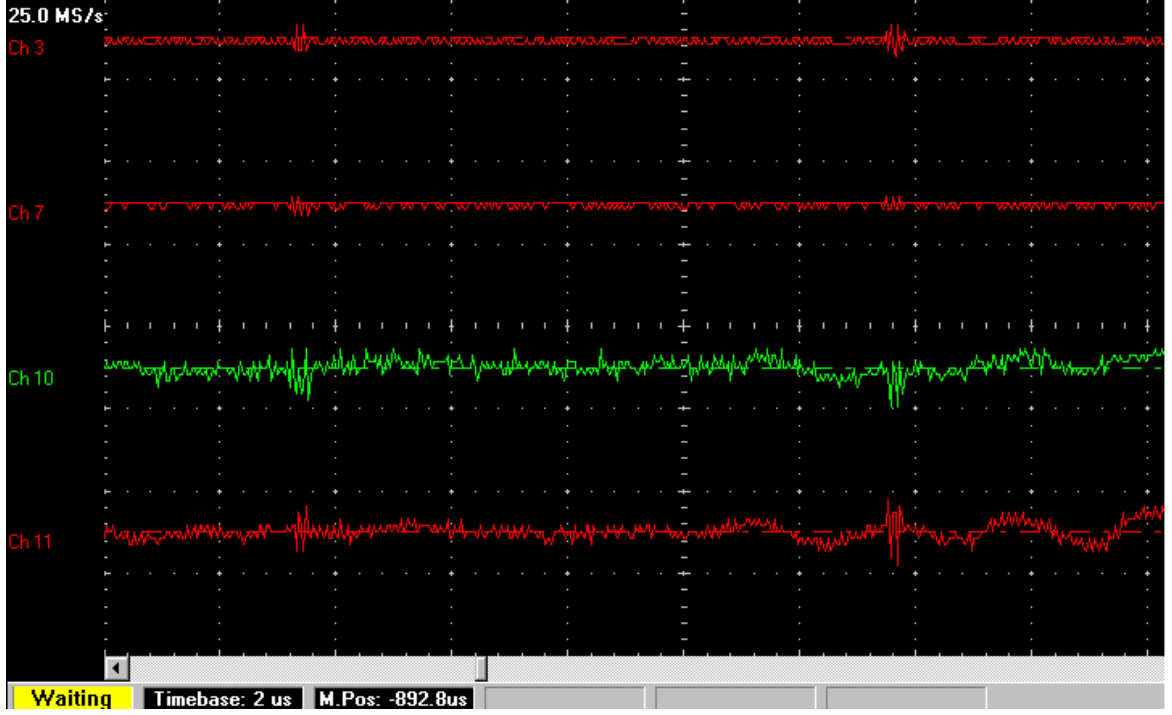


Figure 27. Sampled sequences of 4 channels containing PPM-modulated 6 MHz beatnotes in the presence of atmospheric turbulence.

4.2 Adaptive combining of beatnotes using the LMS algorithm

The discrete complex version of the LMS algorithm can be described by the recursive Eq. (Widrow and Steerns, 1985, Godara, 1997, Benvenuto and Cherubini, 2002):

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e^*(k) \mathbf{r}(k) \quad (4.2.1)$$

The LMS is a recursive algorithm that allows the value of the weight vector \mathbf{w} at time $(k+1)$ to be calculated from the value at time k , using the signal vector at time k . The sampled error signal is obtained from the sampled reference signal and array output, as follows:

$$e(k) = d(k) - \mathbf{w}^H(k) \mathbf{r}(k) \quad (4.2.2)$$

The LMS algorithm described in Eqs. (4.2.1) and (4.2.2) is complex in the sense that the input and output data as well as the weights are all complex values. In the initial simulations and experiments, the target signal used was a constant value, equal to the sum of the average magnitudes of the signals in the signal channels (Compton, 1988). Later on, a different desired signal was utilized as explained later on in the chapter. Since these are early results, obtained prior to the mathematical analysis presented in Chapter 5, the parameter μ was chosen experimentally. However, the unusually large values of μ encountered in this Chapter and in the next, while not normally encountered in the literature, can be explained by the eigenvalue analysis presented in Chapter 5. Here, the focus will be exclusively on the earlier experimental results.

The nominally 6 MHz signal tones were downconverted to complex baseband and input to a least-mean-square algorithm, or LMS. This adaptive algorithm automatically estimates the complex weights required to reconstruct the signal, applies the weights to the complex signal in each channel, and combines the “phased-up” signals in order to maximize power, or SNR. The weights are computed from Eq. (4.2.1) starting with zero initial values. Varying the stepsize, it is possible to control the fraction of the current gradient estimate applied during each update, providing a desired degree of smoothing to the weight estimates. A mathematical discussion of LMS is presented in Chapter 5.

4.2.1 Adaptive combining simulated data: Signal tone and 32-PPM signals

Convergence of the combining weights as a function of sample number has been analyzed. Several cases have been considered, including a signal tone and a PPM signal

observed under ideal conditions. Figure 28 shows the comparison of the convergence of the LMS algorithm for the case of signal tone vs. a 32-PPM signal.

It has been demonstrated that the number of samples required to obtain convergence in the case of M -ary PPM signals with a peak power constraint is M times the number of samples required by a continuous tone, due to decreased total signal energy in the lower duty-cycle modulated waveform. Therefore, for the simulated case of a signal tone, convergence is obtained after 4 samples, while for 32-PPM, 128 samples are required for convergence, as illustrated in Figure 28, for $\mu = 1$ and an introduced phase weight variation of 1 radian between every channel.

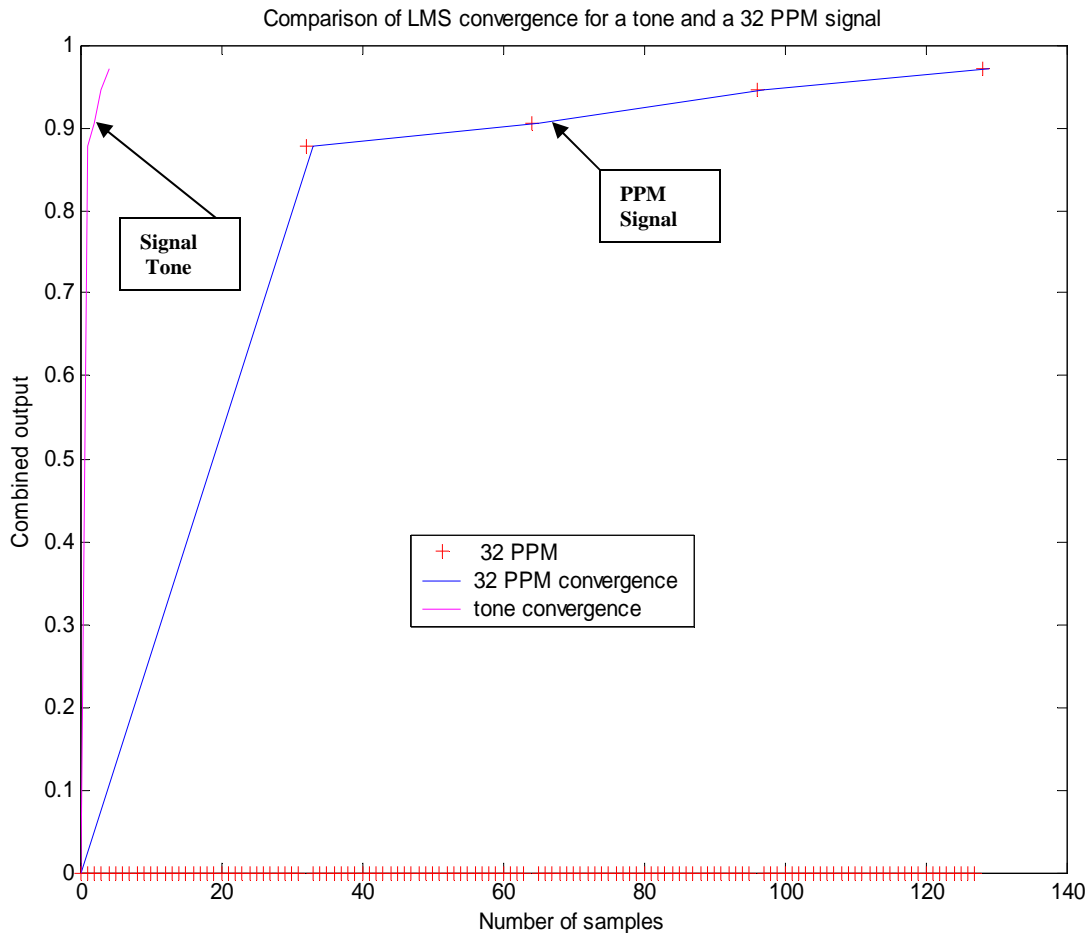


Figure 28. Comparison of output convergence for signal tone and 32-PPM signal (real part of LMS output) .

Figure 29 shows a comparison of convergence for a simulated signal tone and a simulated 32-PPM signal where the stepsize is $\mu = 0.003$ (detector array consists of 16 detectors). Convergence is obtained for a signal tone after 125 samples, and for the 32-PPM signal after 4000 samples, as expected.

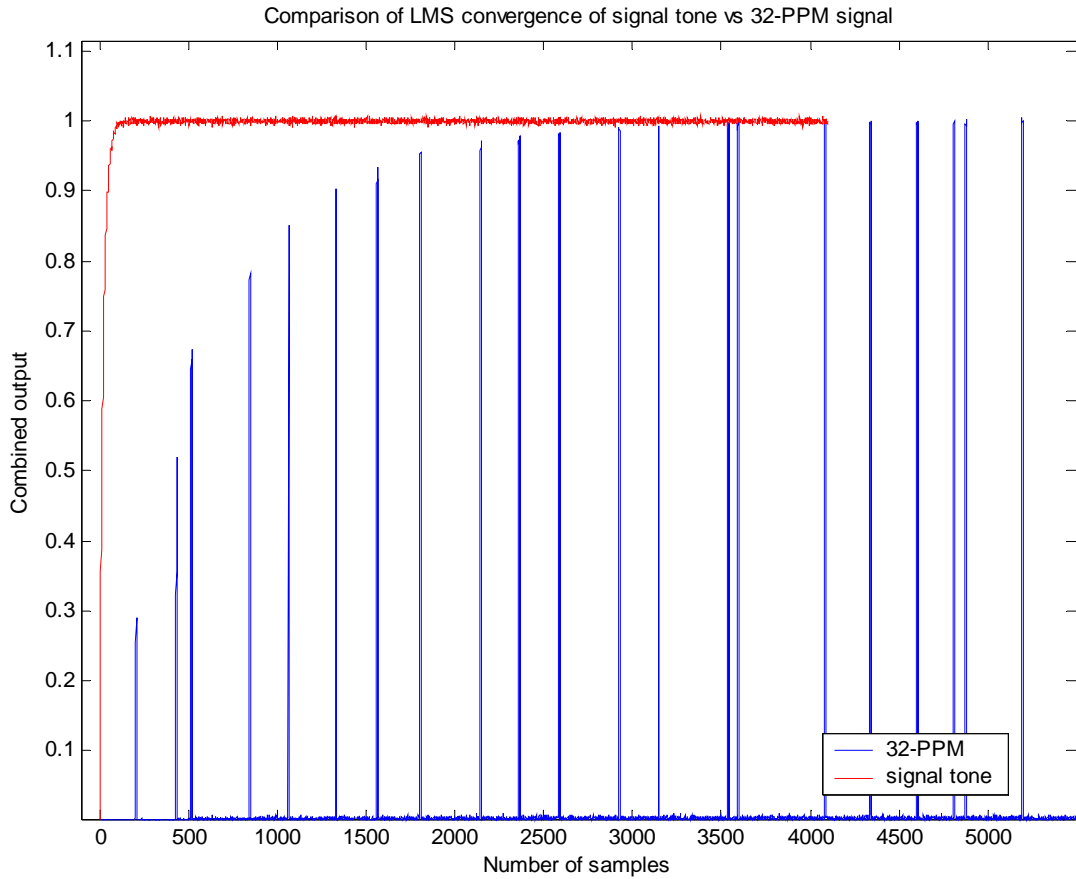


Figure 29. Comparison of convergence for signal tone and a 32-PPM signal with a 16-detector array $\mu = 0.003$.

Other cases have been studied using different values of stepsize. If the stepsize value is increased to $\mu = 0.008$ as shown in Figure 30, the LMS algorithm converges faster, hence only 1000 samples are needed for convergence as opposed to 4000 samples for the previous case.

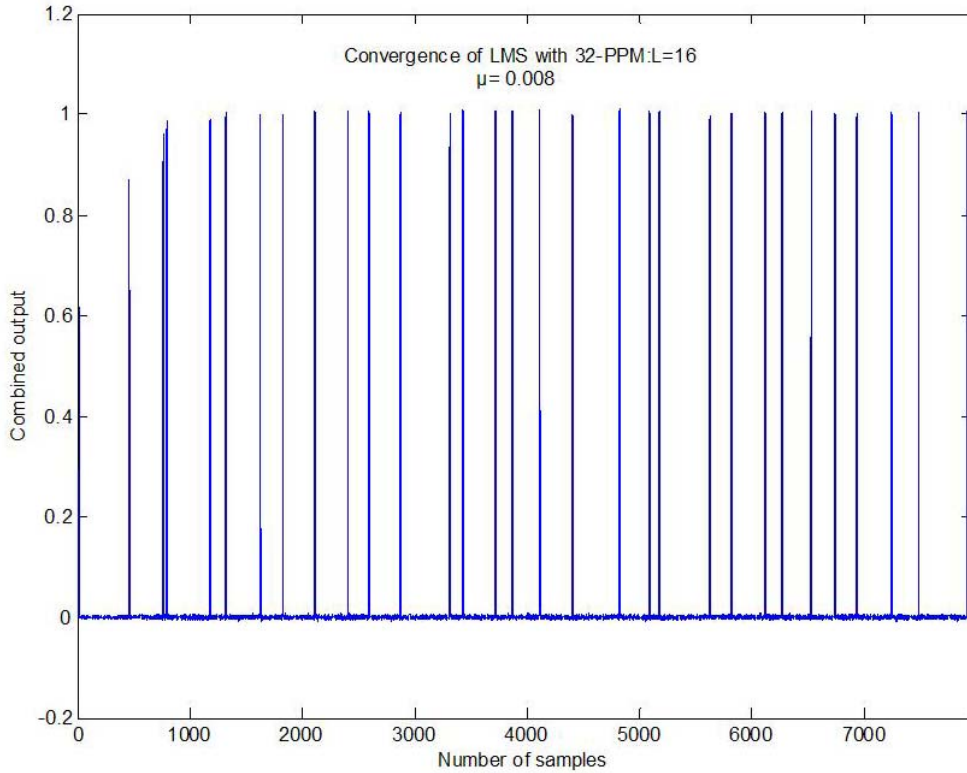


Figure 30. Comparison of convergence for signal tone and a 32-PPM signal with a 16-detector array with $\mu = 0.008$.

4.2.2 Experimental results I: Convergence of LMS algorithm for unmodulated beatnote under ideal conditions and with atmospheric turbulence

In these initial experiments, the reference signal is a constant value calculated based on the addition of the magnitude of the signal in the four channels, which resulted in a value of 0.06. Therefore, the error signal obtained is a complex number that contains the phase information required. The error signal has to be minimized in order for the phase weights to converge; at that point, the four signals will be phased up and the combined output will be maximized. The weights are computed from Eq. (4.2.1) starting with zero initial values. Varying the stepsize, it is possible to control what fraction of the latest weight

estimate is applied to the current weight during each update, providing additional smoothing to the weight estimates.

Referring to Figures 31, 32, and 33 with $\mu = 10$, the stepsize is so small that the LMS algorithm cannot keep up with the phase variations in the beatnote. The combined output signal shows that the four channels are not perfectly combined as it oscillates and never reaches its maximum value of 0.06. The weights have a sawtooth shape, which is due to continuously changing phase in the downconverted output, which is not exactly at zero frequency, but very close to it. Note that the error signal never settles down to a small value in Figure 33.

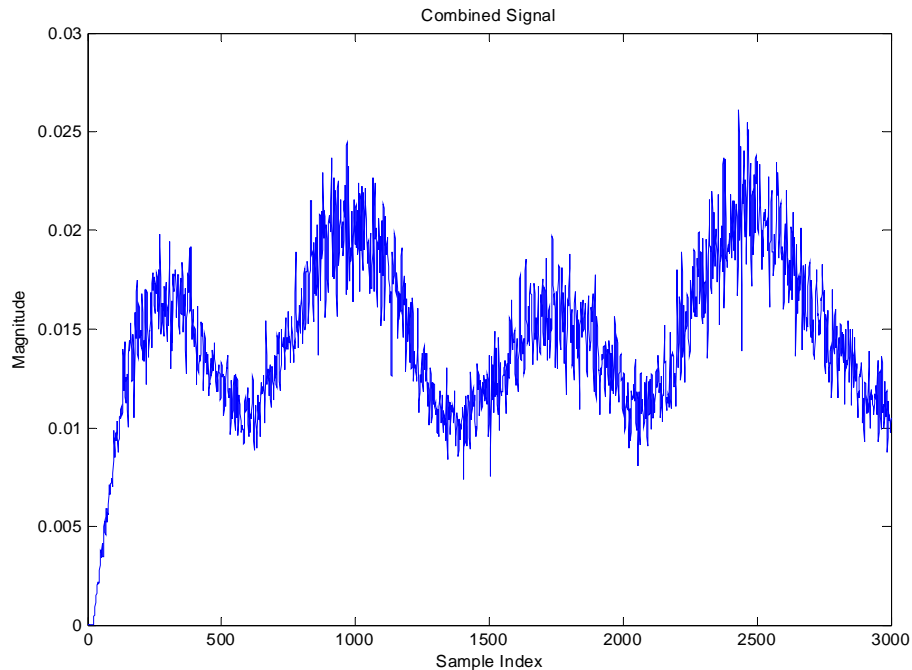


Figure 31. Combined output signal with $\mu = 10$.

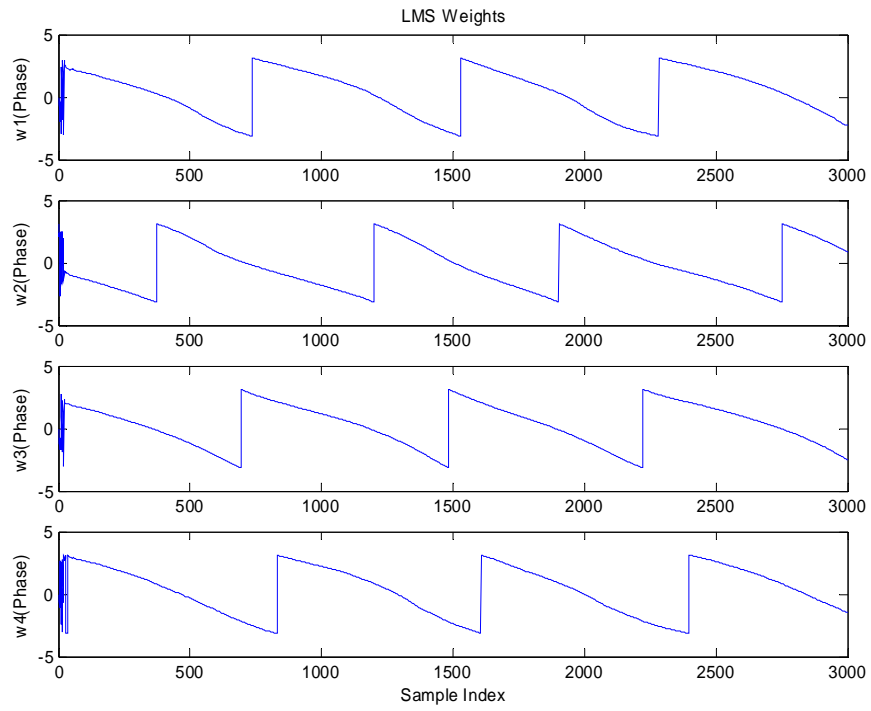


Figure 32. Phase of the weights with $\mu = 10$.

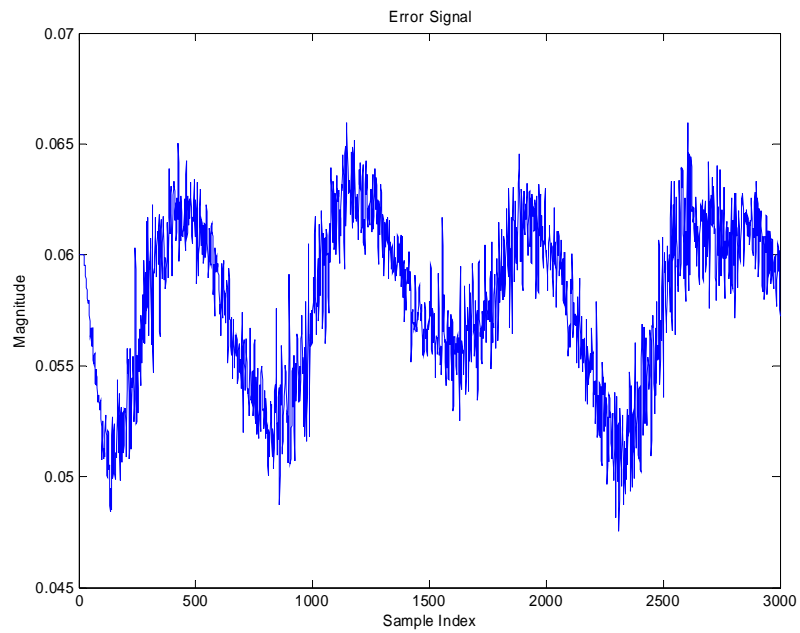


Figure 33. Error signal for $\mu = 10$.

Referring to Figures 34, 35, and 36, we observe that as we increase the value of the stepsize, now set to $\mu = 100$, we get greatly improved performance. The combined output shown in Figure 34, has increased in value approaching its maximum. The error signal has decreased, showing partial convergence of the weights as we can see in Figure 36.

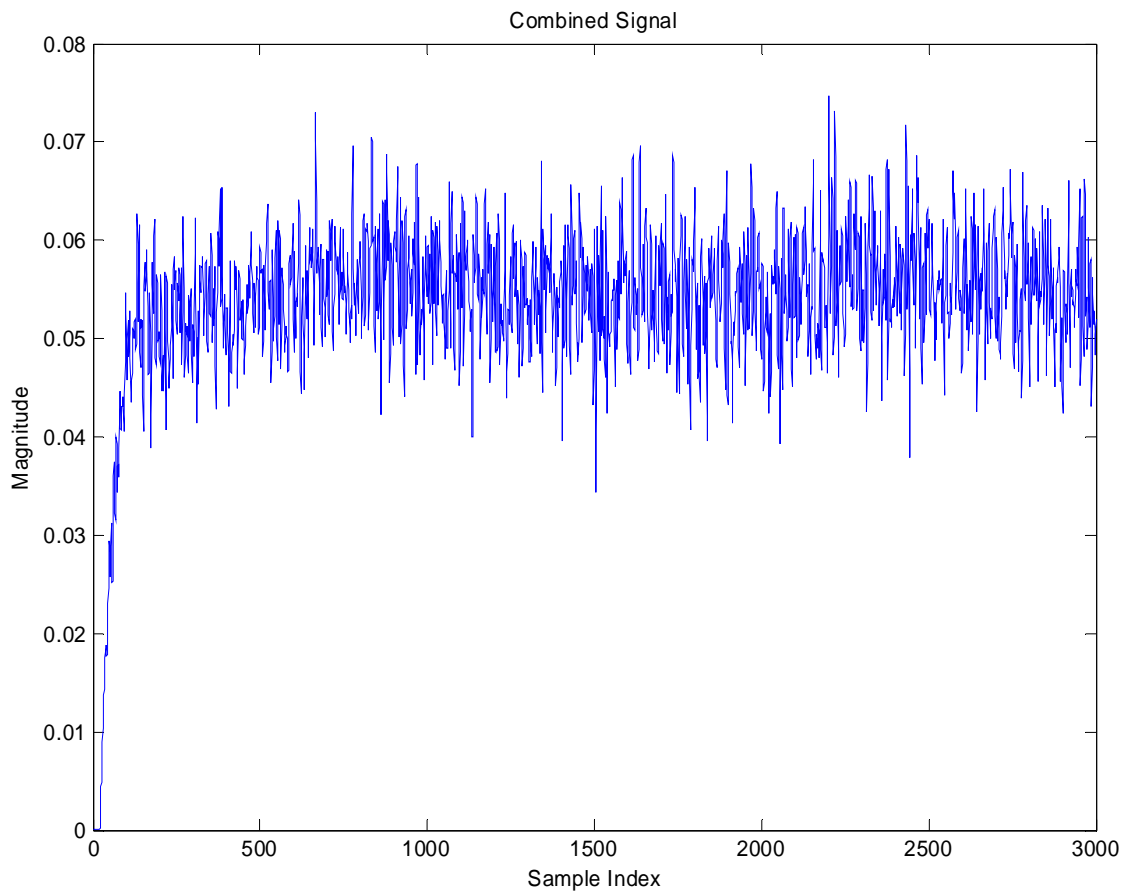


Figure 34. Combined output power with $\mu = 100$.

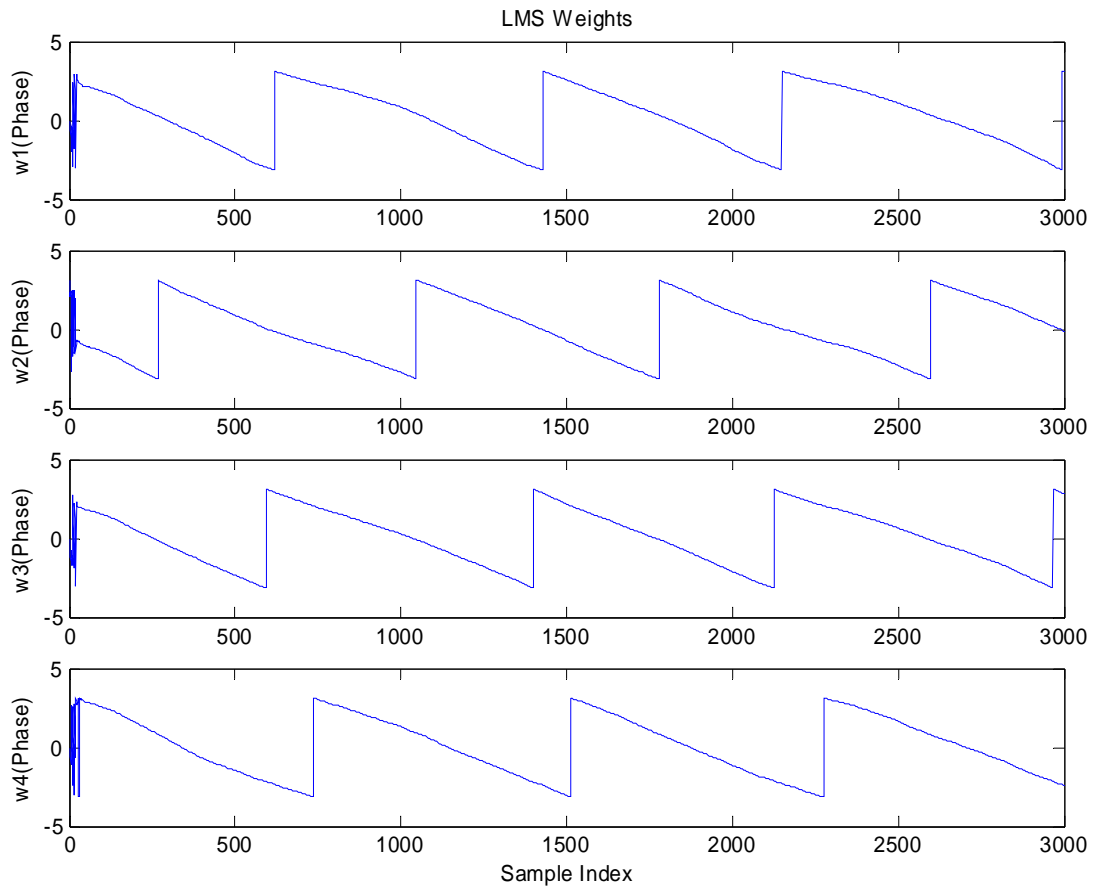


Figure 35. Phase of the weights with $\mu = 100$.

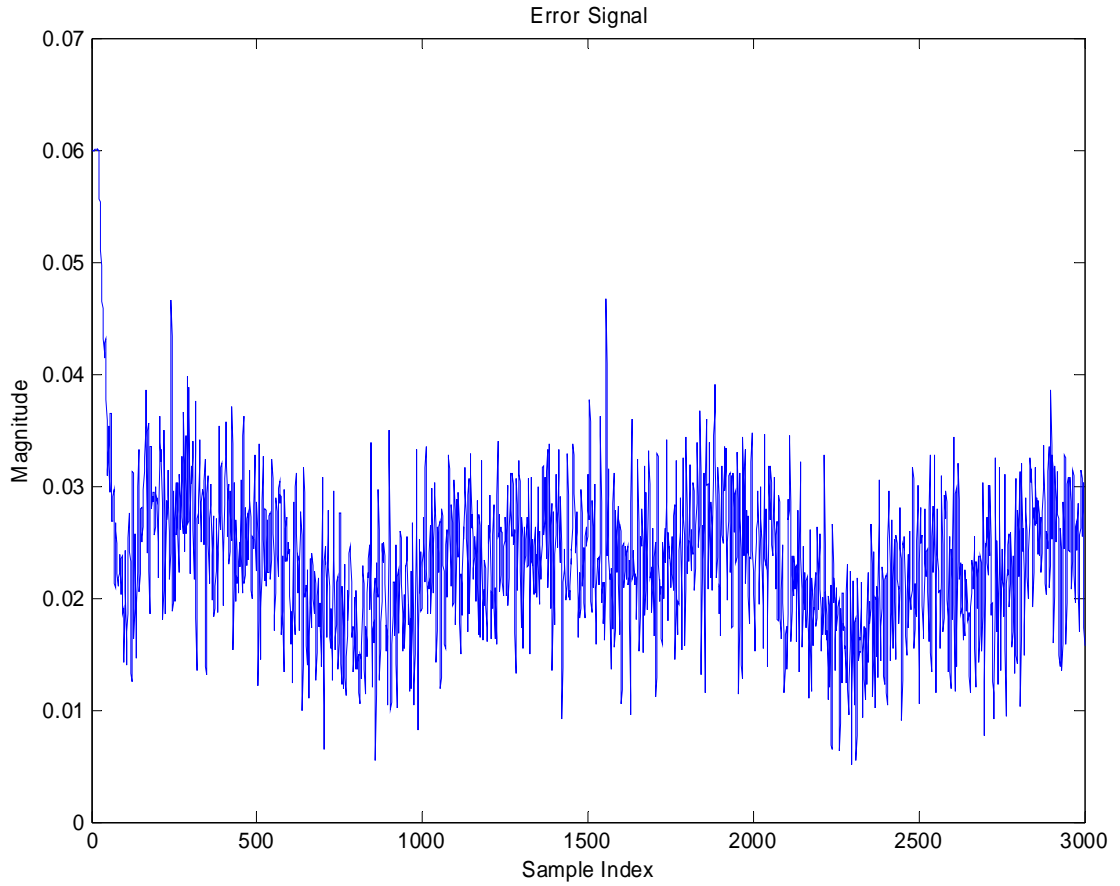


Figure 36. Error signal with $\mu = 100$.

Finally, when the stepsize is large enough so that the LMS algorithm is able to keep up with the phase rotation of the complex downconverted beatnote, at $\mu = 1000$, the combined output signal reached its expected maximum value of 0.06. With this optimum value of μ , the error approaches zero and it is concluded that the signals are phased up. Figures 37, 38, and 39 illustrate the combined output, phase of the weights, and error signal respectively.

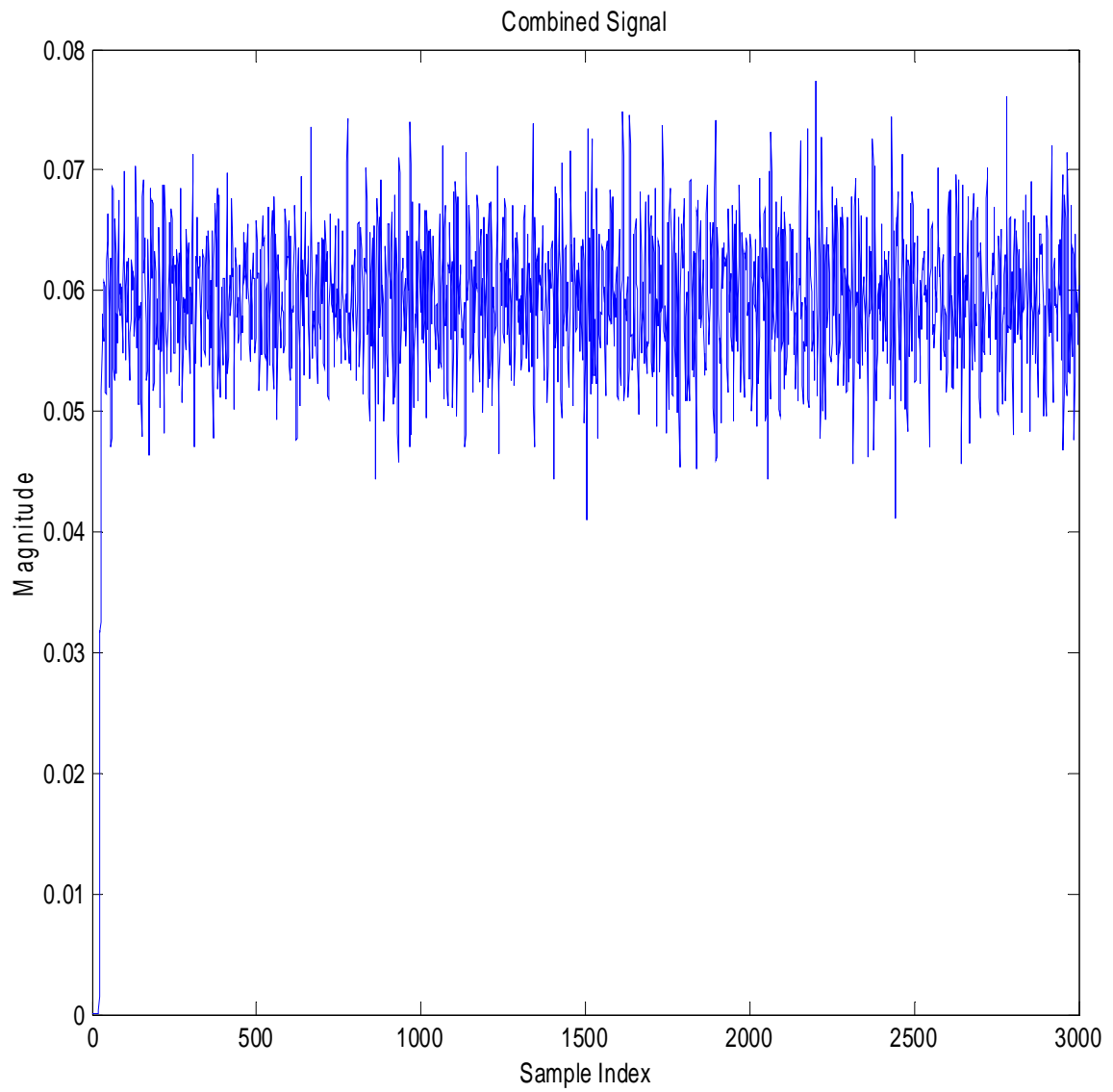


Figure 37. Combined output signal with $\mu = 1000$.

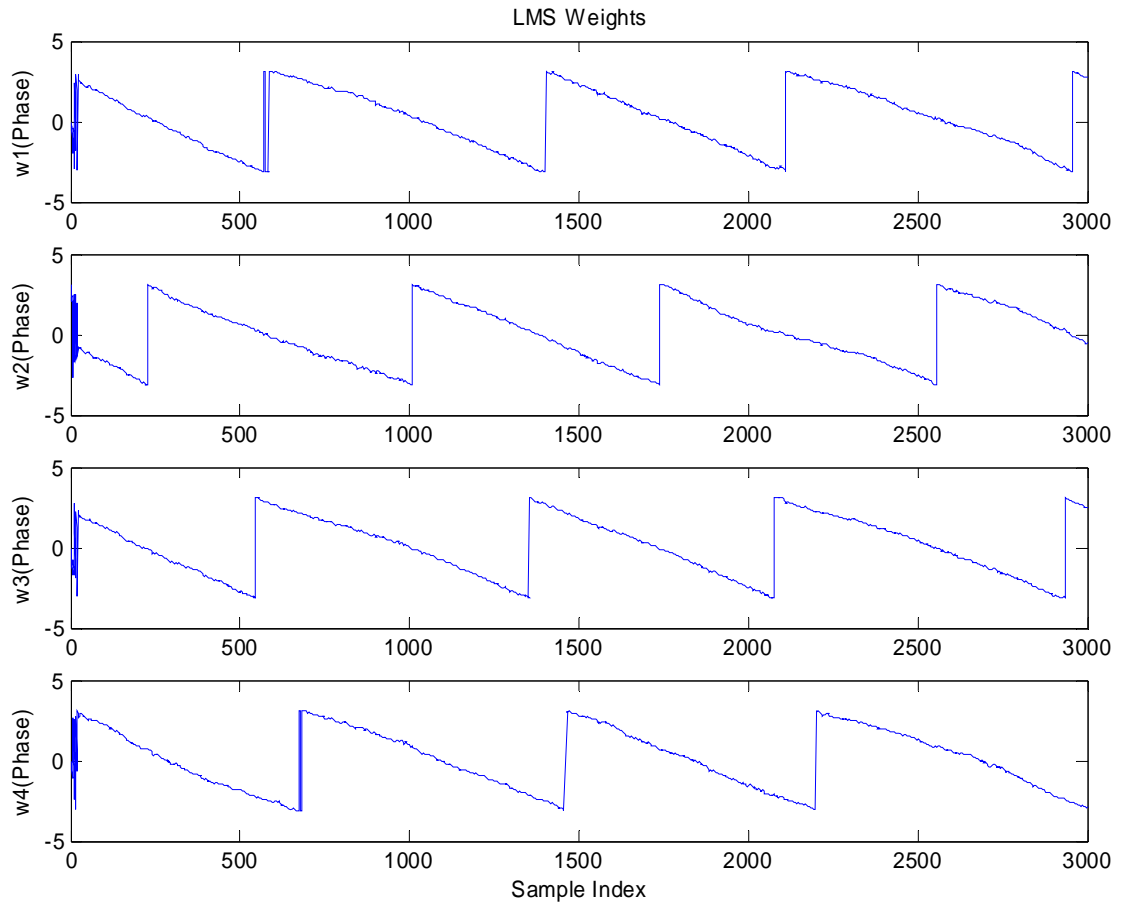


Figure 38. Phase of the weights of the four different channels with $\mu = 1000$.

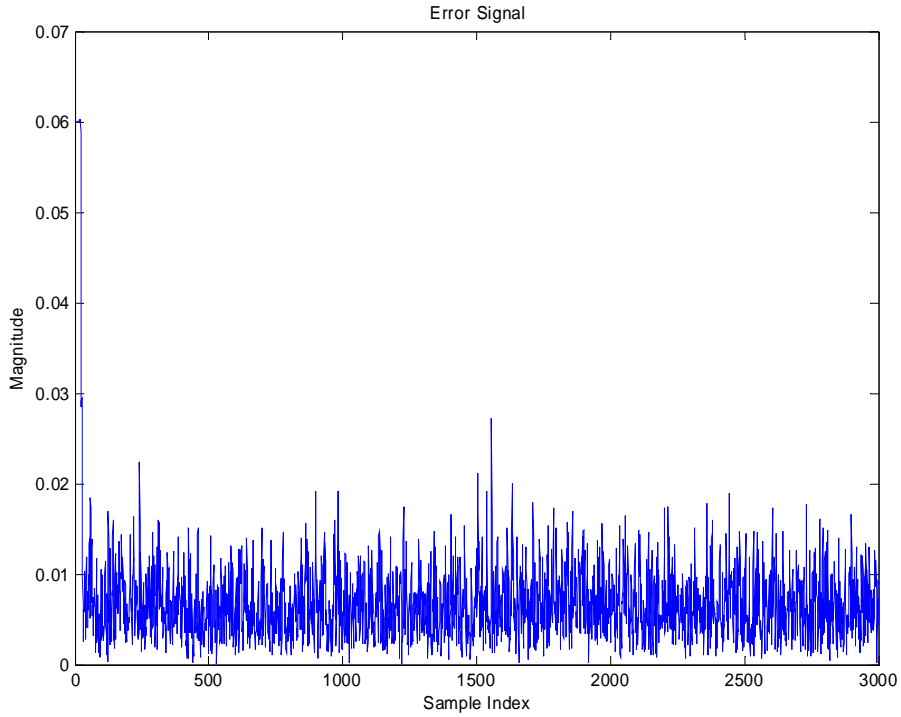


Figure 39. Error signal with $\mu = 1000$.

These results illustrate that increasing the stepsize allows the LMS algorithm to follow and track the phase rotation of the complex downconverted beatnote. It was found that a good value of μ for the particular case discussed here is 1000. For this value there is accurate tracking of the signals and accordingly, the error signal approaches zero, and maximum combined output is achieved.

It is important to note that usually in books and papers, the stepsize is shown to be much smaller than one, but that is because the signal is assumed to be of unity amplitude. In this experiment, the signal levels that we are dealing with are very small since there is not enough amplification after detection. Therefore, large values of stepsize are needed to provide adequate updates to the weights.

The purpose is to analyze the case of a signal tone received in the presence of simulated atmospheric turbulence conditions using the rotating plexiglass plate shown in Figure 20. The reference signal used in the algorithm for this situation resulted in a value of 0.0036. As in our previous case, four channels that contained significant signal were identified, and at a certain time 104,128 samples were synchronously collected from each channel. After some experimentation, it was determined that for this data set good results could be obtained by correlating over 10,000 samples, and using a stepsize of 20,000 as the signal is even weaker than for the ideal case due to the loss introduced by the atmospheric turbulence added to the system. Figure 40 shows the combined output that reaches its maximum value of 0.0036. Convergence of the LMS algorithm is accomplished after 200 samples, and therefore the acquisition time is 8 μ s.

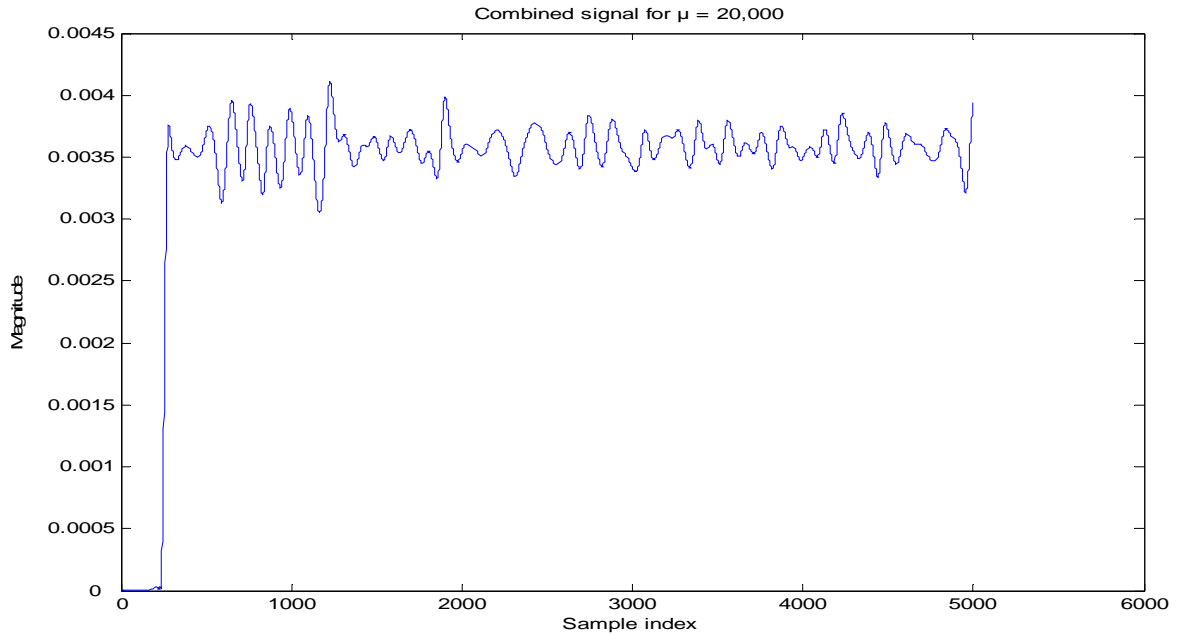


Figure 40. Combined output power of the beatnote signal in the presence of simulated atmospheric turbulence with $\mu = 20$.

In order to minimize higher frequency noise contributions in this case, we use a narrower filter bandwidth on the signal processing block.

Figure 41 shows the phase of the weights with $\mu = 20,000$. For this case, the weights also have a sawtooth shape, due to the continuously changing phase in the downconverted output, as it is not exactly at 0 Hz.

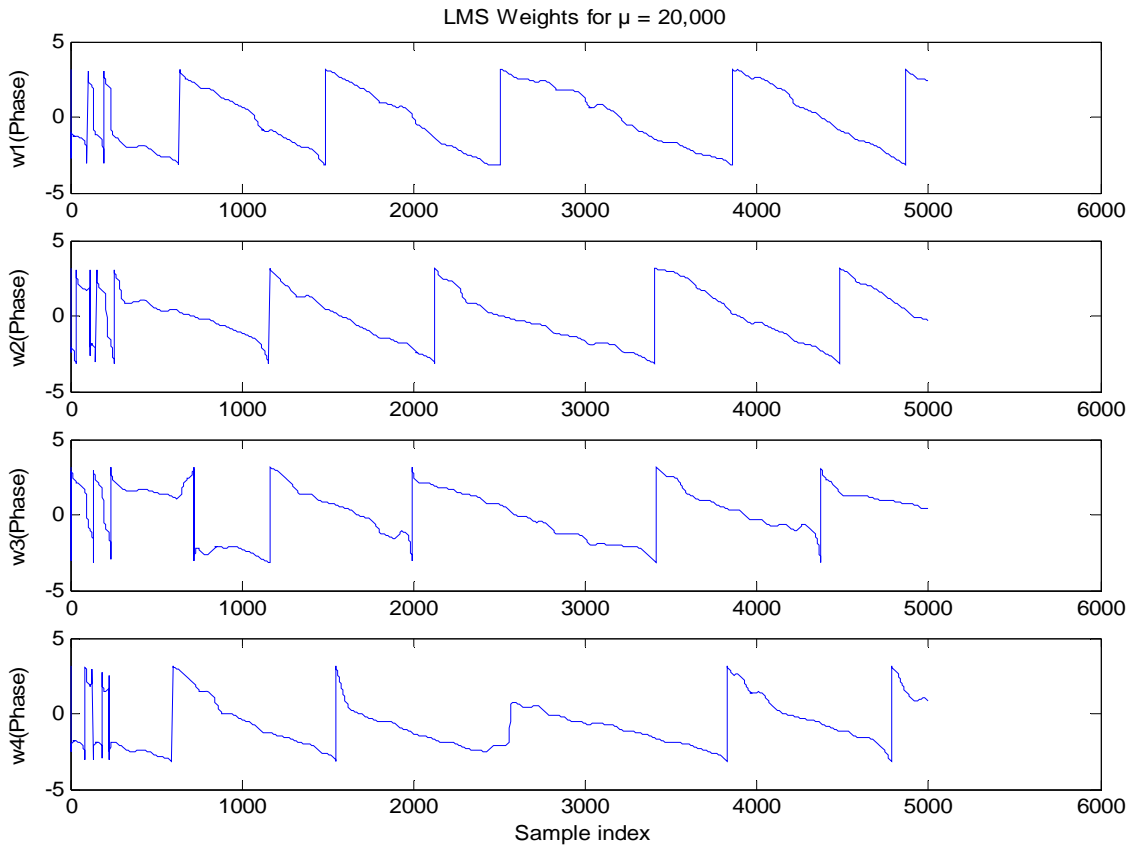


Figure 41. Phase of the weights with $\mu = 20,000$.

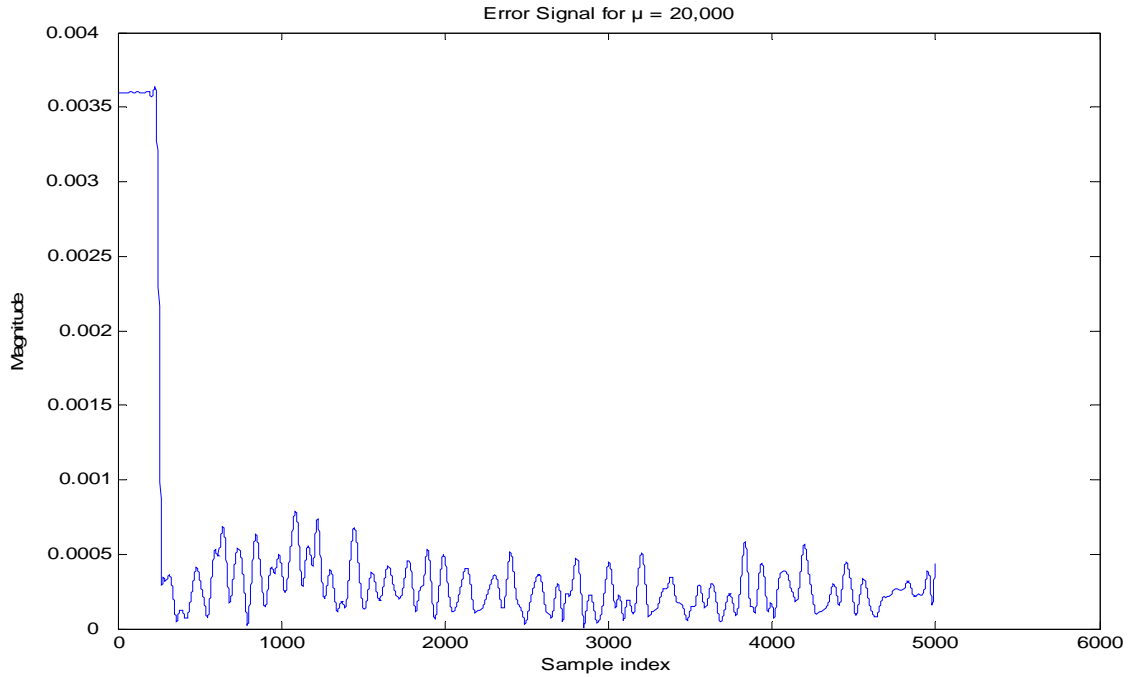


Figure 42. Error signal with $\mu = 20,000$.

Figure 42 shows that the error signal approaches zero when the stepsize is $\mu = 20,000$. Therefore, convergence of the LMS weights has been achieved and the combined output value has been maximized with minimum error.

These results illustrate, that increasing the stepsize allows the LMS algorithm to follow and track the phase rotation of the complex downconverted beatnote in the presence of atmospheric turbulence conditions. It was found that a good value of μ for the particular case discussed here is 20,000 when there is accurate tracking of the signals and accordingly, the error signal approaches zero, and maximum combined output is achieved. It is important to note that usually books and papers, the stepsize is shown to be much smaller than one, but that is because the signal is assumed to be of unity amplitude.

In our experiment, the signal levels that we are dealing with are very small as there is not enough amplification after detection. Therefore, large values of stepsize are needed to provide adequate updates to the weights.

4.2.3 Experimental results II: Convergence of the LMS algorithm with PPM signals and no atmospheric turbulence

We first consider the case using a stepsize of $\mu = 1$ (considered to be small for the experimentally recorded data, which are on the order of 0.01 for the individual channels). The sum of the magnitudes of the signals in the four selected channels is approximately 0.186. Figure 43 shows the combined output (and weighted channel components) of the LMS combiner for this case; with a stepsize of 1, the LMS algorithm cannot keep up with the phase variations in the beatnote, hence the combined output signal never reaches its maximum value of 0.186; instead, it reaches only about 0.037.

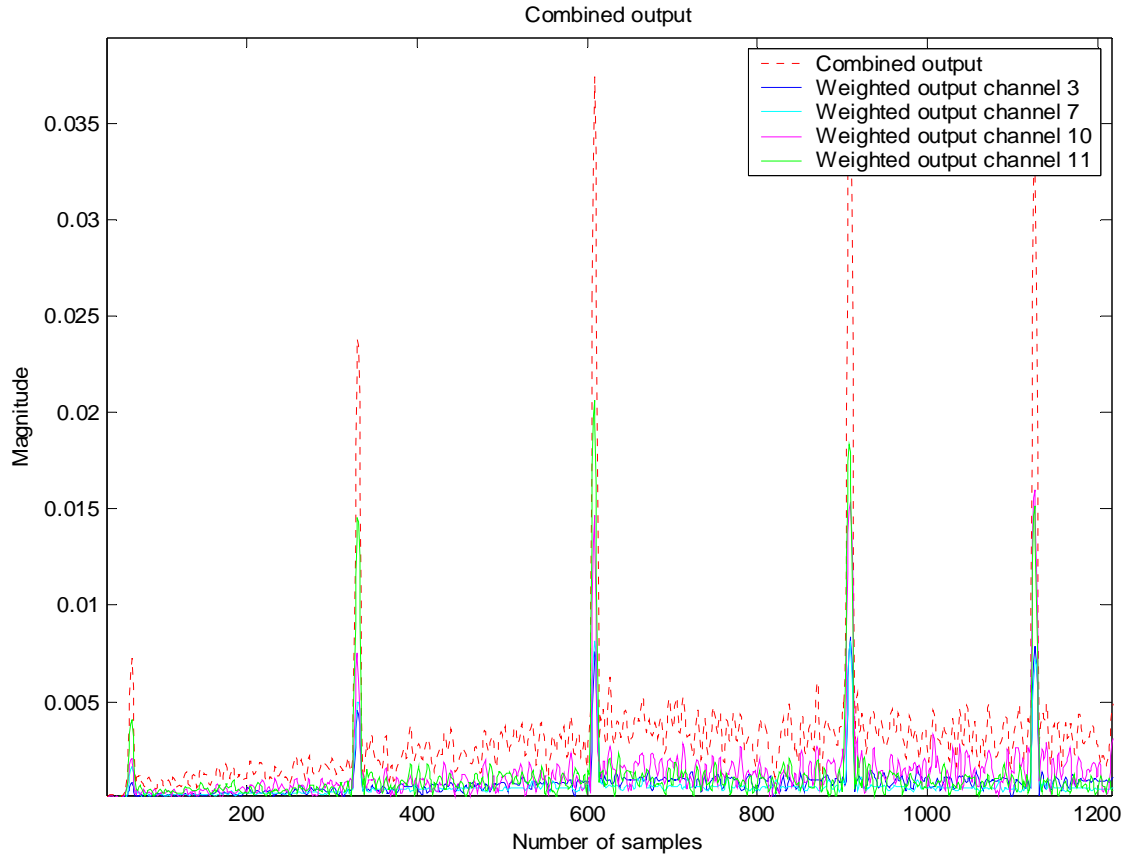


Figure 43. Combined output with $\mu = 1$.

When a larger stepsize is used, $\mu = 7$, the combined output achieves the desired value of 0.186 as illustrated in Figure 44. We see that the combined output reaches its desired maximum value after approximately 800 samples; this translates to an acquisition time of approximately $32 \mu\text{s}$.

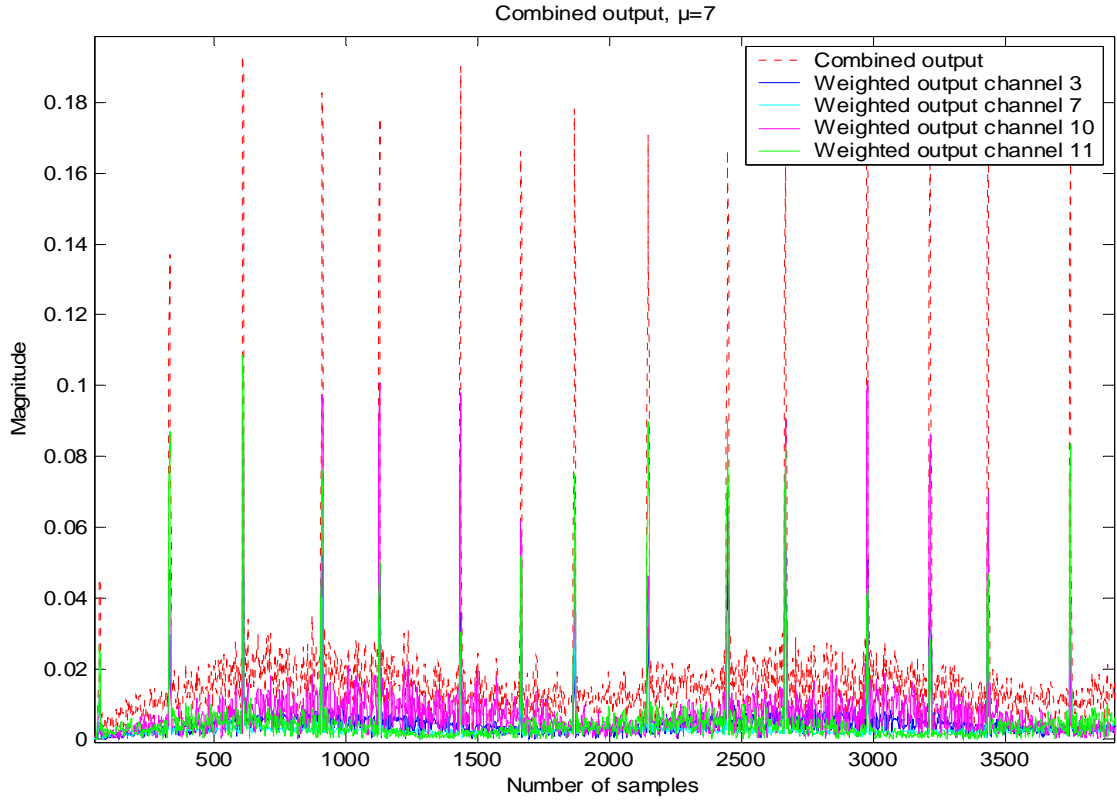


Figure 44. Combined output with $\mu = 7$.

Figure 45 shows the behavior of the phase of the combining weights, as a function of time (or samples). We observe that the phase of the weights has a sawtooth shape due to the continuously changing phase in the downconverted output, which is not exactly at zero frequency.

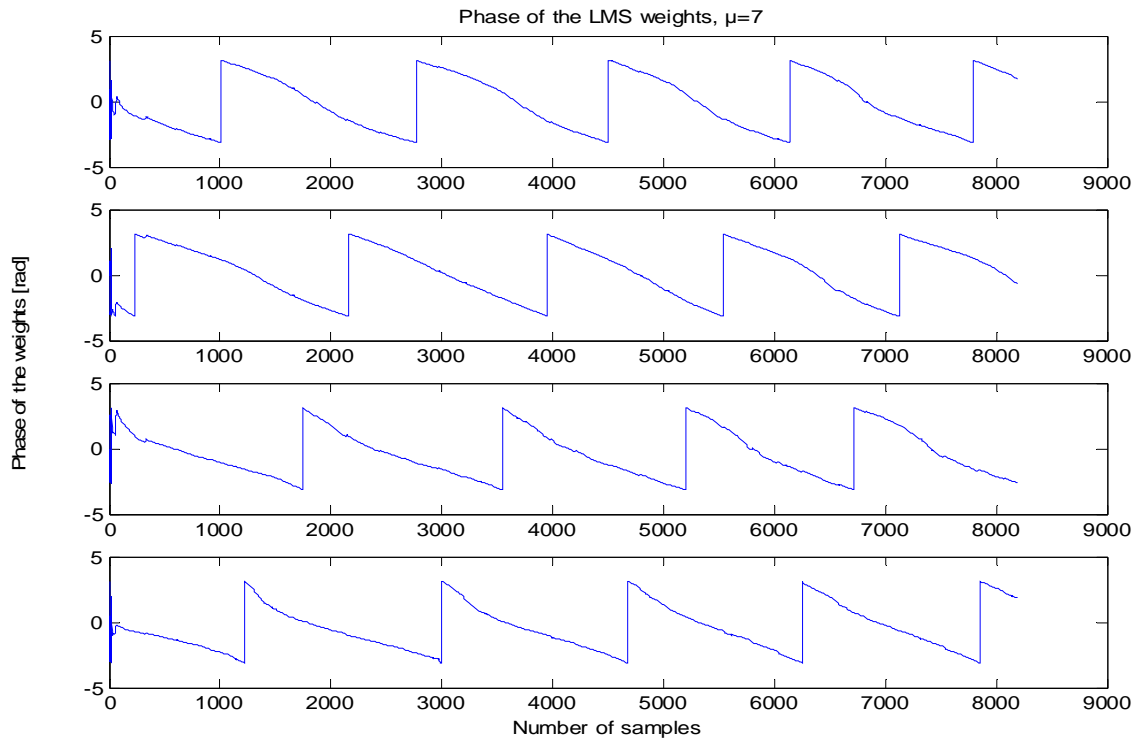


Figure 45. Phase of the weights for $\mu = 7$.

Figure 46 also shows an individual combined pulse and its weighted components in greater detail. The addition of the magnitudes of the four channels is 0.186; indeed, the components sum to the expected value, verifying the validity of the instantaneous combining operation.

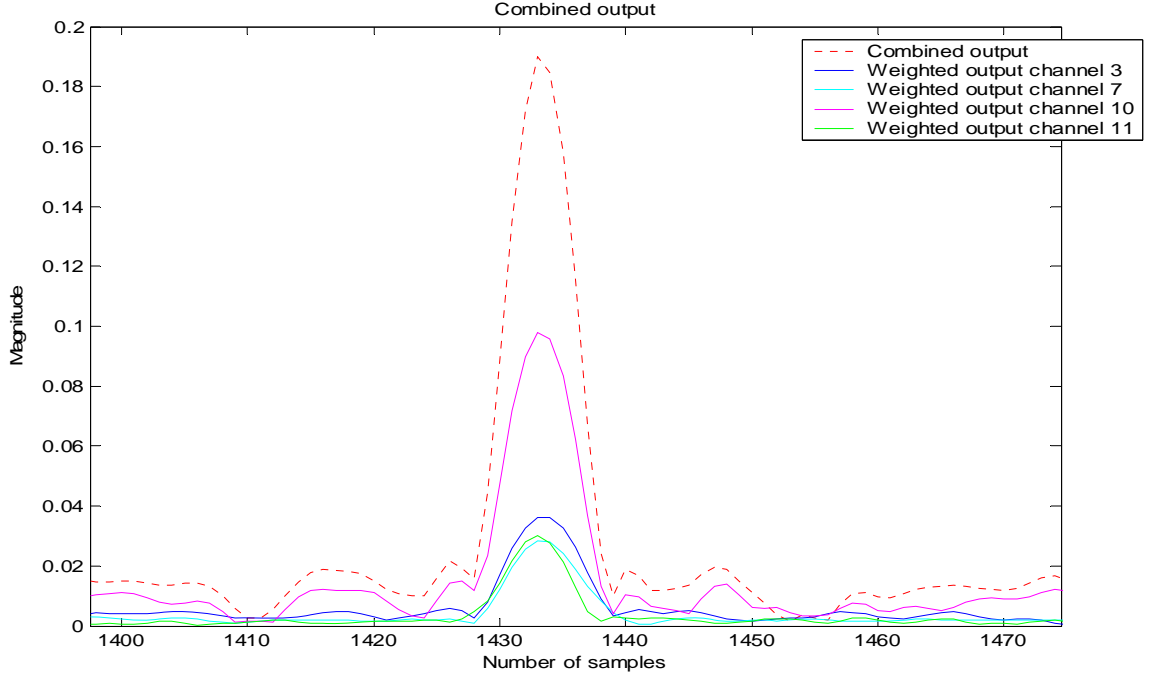


Figure 46. Combined output and weighted signal components with $\mu = 7$.

In summary, it can be seen that the experimentally obtained PPM signals are combined correctly with the larger stepsize, and converge to their desired final value in less than a millisecond.

4.2.4 Convergence of LMS algorithm in the presence of spatial distortions caused by a static plexiglass plate in the optical path

For the case of combining detector array output signals spatially distorted by the plexiglass plate, the desired signal magnitude is the addition of the average magnitudes of the individual channels, which in this case turned out to be 0.063. Initially, we attempt to combine adaptively using a stepsize of 8, however it can be seen from Figure 47 that the stepsize is too small, hence the LMS algorithm cannot keep up with the residual phase variations and only attains a magnitude of 0.033.

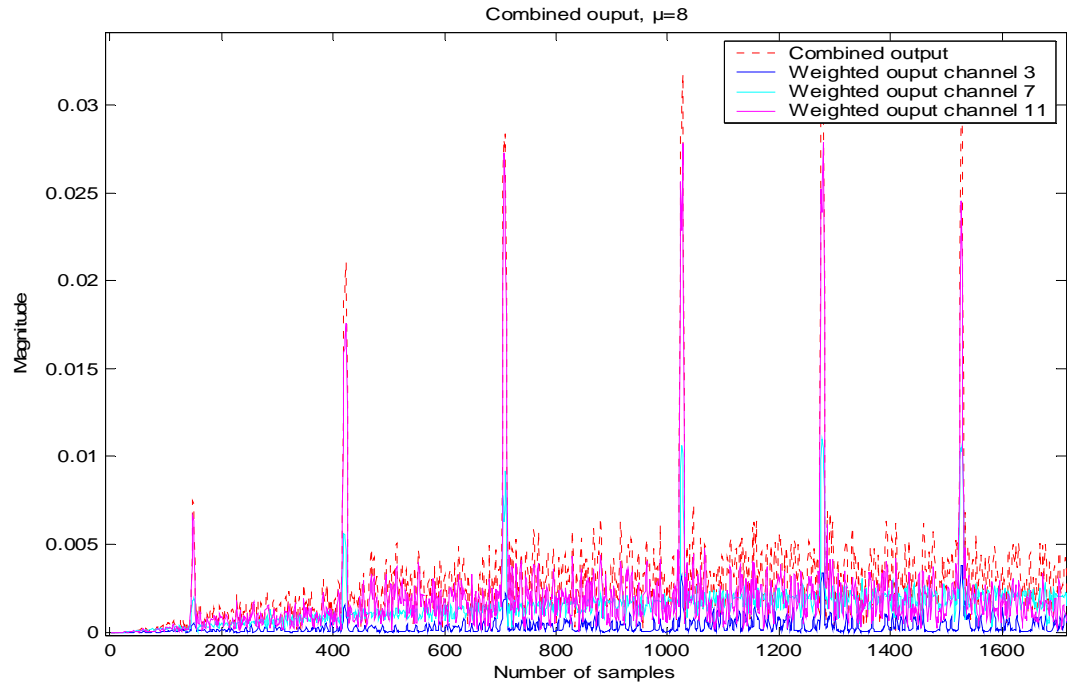


Figure 47. Combined output with $\mu = 8$.

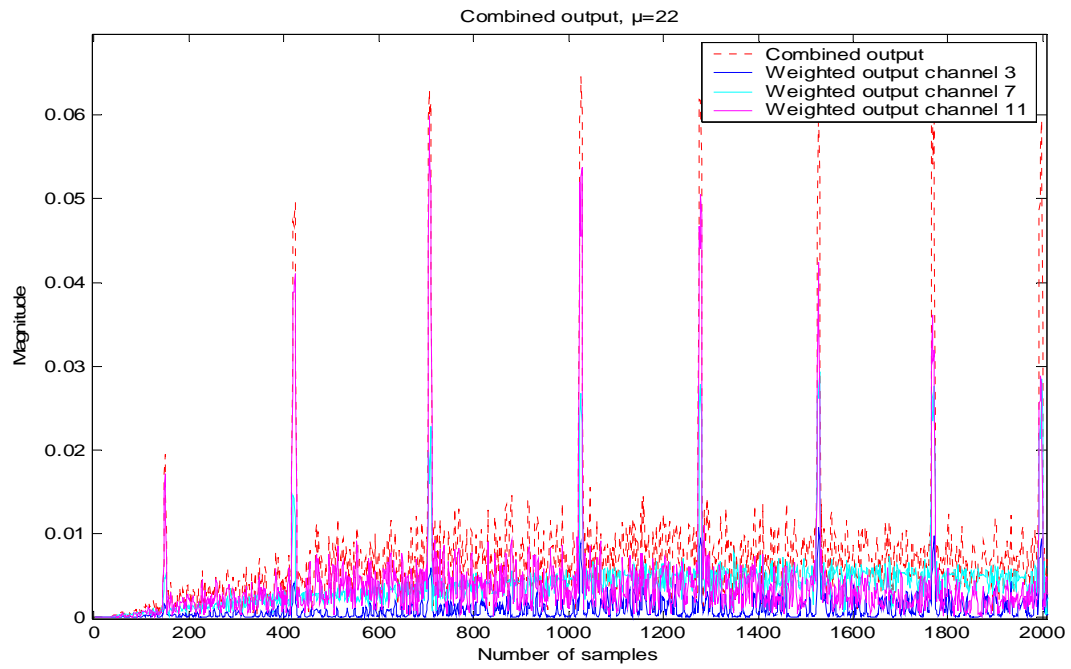


Figure 48. Combined output with $\mu = 22$.

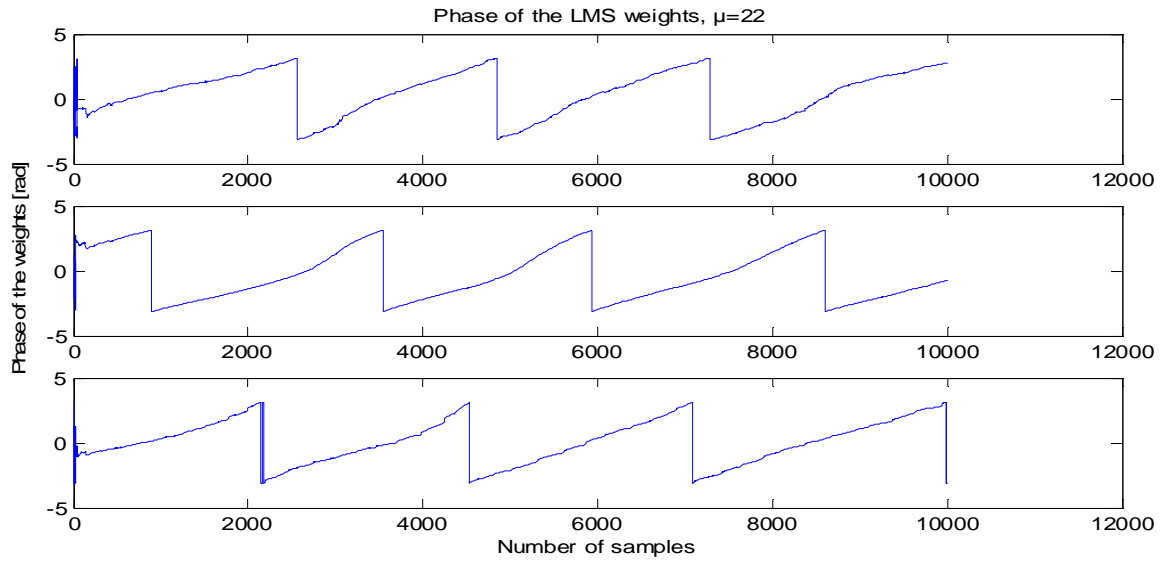


Figure 49. Phases of the weights for $\mu = 22$.

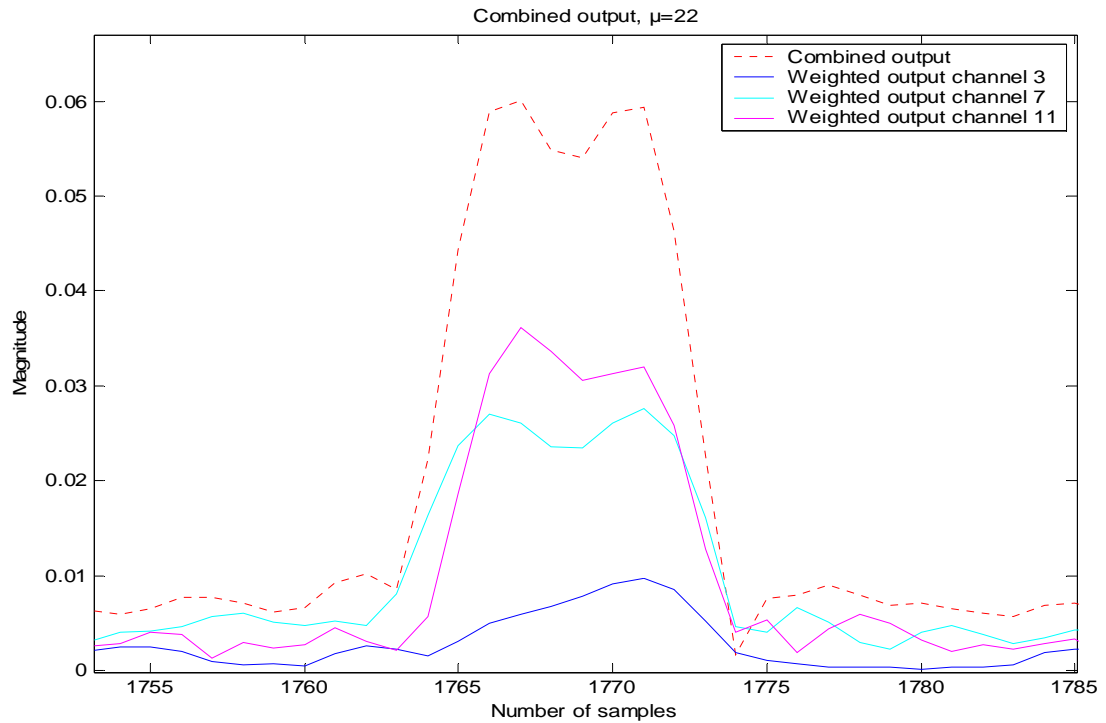


Figure 50. Combined output with $\mu = 22$.

Referring to Figures 48, 49, and 50, we observe that as we increase the value of the stepsize to $\mu = 22$, we get greatly improved combining performance. The combined output shown in Figure 48, and in more detail in Figure 50, has increased, approaching its maximum value of 0.063. At this point, the stepsize is large enough so that the LMS algorithm is able to keep up with the phase rotation of the complex downconverted beatnote. These results illustrate, that increasing the stepsize allows the LMS algorithm to follow and track the phase rotation of the complex downconverted beatnote. For this value there is accurate tracking of the signals and accordingly, the error signal approaches zero, and maximum combined output is achieved.

Chapter 5: The LMS Algorithm and Its Application to Coherent Optical Signal Reception

In this system, the well-known LMS algorithm is used for the purposes of combining the outputs of the individual FPA channels in order to generate a combined channel whose performance is close to that of a single channel without any signal losses due to atmospherically induced phase distortions. The LMS algorithm was briefly introduced in Chapter 4, and this Chapter discusses it in greater depth.

In Section 5.1, the signal combining problem and the goal of minimizing the mean squared error are presented along with the optimal (in the minimum mean squared error sense) Wiener solution. This discussion leads into a discussion of the LMS algorithm for stationary, ergodic signals in Section 5.2. The time adaptive nature of LMS enables it to work well in many cases even in non-stationary environments (Sayed, 2003), and some of the mathematical implications of non-stationarity for the case of reception of signals with an FPA are considered in Section 5.3. Further discussion of the effects of a leakage signal (defined in Section 5.3) are presented in Section 5.4. A variant of the LMS algorithm presented in Compton, (1988) that utilizes a constant target signal is considered in Section 5.5, and the convergence time performance of the LMS algorithm with real world data is considered in Section 5.6.

5.1 The signal model

The signal is modeled as a desired component corrupted by Gaussian noise. Let $d(k)$ denote the desired communications signal. Let $\mathbf{s}(k)$ denote the signal vector, which contains information on the desired signal $d(k)$, and let $\mathbf{n}(k)$ denote a zero-mean vector additive white Gaussian noise (AWGN) process. All signals are represented in complex baseband. The vector AWGN process $\mathbf{n}(k)$ is assumed to satisfy:

$$E[\mathbf{n}(i)\mathbf{n}(j)^H] = \begin{cases} 2\sigma^2\mathbf{I} & i = j \\ \mathbf{0} & i \neq j \end{cases} \quad (5.1.1)$$

where σ^2 is the variance of both the real and imaginary parts of the Gaussian noise in each of the individual channels of the FPA, yielding a net variance of $2\sigma^2$ per complex baseband channel.

The received vector signal at time k can be written as:

$$\mathbf{r}(k) = \mathbf{s}(k) + \mathbf{n}(k) \quad (5.1.2)$$

where, in the FPA case:

$$\mathbf{s}(k) = \mathbf{a}d(k) \quad (5.1.3)$$

Here \mathbf{a} is the vector of complex magnitudes and phases at the focal plane array (FPA) detector elements due to the focal plane field distribution and $d(k)$ is the desired communications signal.

The output of the adaptive filter is given by:

$$y(k) = \mathbf{w}^H(k)\mathbf{r}(k) \quad (5.1.4)$$

and the goal is to minimize the mean squared error, which is defined as:

$$\text{MSE} = E \left[\left| d(k) - y(k) \right|^2 \right] \quad (5.1.5)$$

The goal of the LMS algorithm is to find the optimal weight vector \mathbf{w}_{opt} such that the MSE defined in Eq. (5.1.5) is minimized. However, an alternative to minimizing the MSE is minimizing the BER as shown by Chen et al. (2001) since the MSE criterion may not always be optimal from a BER standpoint. This subject is not considered further in this thesis, but is a good direction for future research.

Following the treatments in both Haykin, (1991) and Sayed, (2003), and assuming ergodicity and wide-sense stationarity, define the autocorrelation and cross-correlation matrices shown below:

$$\mathbf{R} = E \left[\mathbf{r}(k) \mathbf{r}^H(k) \right] \quad (5.1.6)$$

$$\mathbf{p} = E \left[\mathbf{r}(k) d^*(k) \right] \quad (5.1.7)$$

The optimal solution, or Wiener solution, is given by:

$$\mathbf{w}_{opt} = \mathbf{R}^{-1} \mathbf{p} \quad (5.1.8)$$

If the input stochastic process is both stationary and ergodic, then Eq. (5.1.8) yields the optimal weight vector. If the process is non-stationary, but if both $\mathbf{R}(k)$ and $\mathbf{p}(k)$ are perfectly known as a function of time, then minimization of the expected value of the mean squared error in Eq. (5.1.5) would yield:

$$\mathbf{w}_{opt}(k) = \mathbf{R}^{-1}(k) \mathbf{p}(k) \quad (5.1.9)$$

5.2 Gradient descent and LMS

Theoretical analyses of the behavior of the LMS algorithm are widely available in the literature, and the results from Haykin, (1991) will be stated here. The LMS algorithm is designed to perform a gradient descent approximation that converges closely to the Wiener solution. The treatment of Haykin, (1991) assumes that the input and the desired signal are jointly wide-sense stationary, but this assumption is not true in the system under consideration. The implications of the non-stationarity will be briefly considered and are a topic of future research although the LMS algorithm is both widely and successfully used for the tracking of non-stationary signals (Sayed, 2003).

Consider the problem of adaptively finding a good estimate of the weights in real time. In that case, the gradient of the MSE with respect to the weight vector $\mathbf{w}(k)$ is given by (Haykin, 1996):

$$\nabla_{\mathbf{w}(k)} \text{MSE}(k) = -2\mathbf{p}(k) + 2\mathbf{R}(k)\mathbf{w}(k) \quad (5.2.1)$$

It should be noted that substitution of the Wiener solution from Eq. (5.1.8) into the gradient in Eq. (5.2.1) results in a gradient of zero as expected.

The statistical quantities \mathbf{p} and \mathbf{R} are usually difficult to obtain. The gradient in Eq. (5.2.1) can be approximated using the instantaneous sample $\mathbf{r}(k)$ and reference $d(k)$ as shown below:

$$\begin{aligned}
\nabla_{\mathbf{w}(k)} \text{MSE}(k) &= -2\mathbf{p}(k) + 2\mathbf{R}(k)\mathbf{w}(k) \\
&= -2E[\mathbf{r}(k)d^*(k)] + 2E[\mathbf{r}(k)\mathbf{r}^H(k)]\mathbf{w}(k) \\
&\approx -2\mathbf{r}(k)(d(k) - \mathbf{w}^H(k)\mathbf{r}(k))^* \\
&= -2\mathbf{r}(k)e^*(k)
\end{aligned} \tag{5.2.2}$$

where $e(k) = d(k) - \mathbf{w}^H(k)\mathbf{r}(k)$. The LMS approximation to gradient descent is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e^*(k)\mathbf{r}(k) \tag{5.2.3}$$

The learning rate parameter μ , which includes the factor of 2 from Eq. (5.2.2), must be chosen large enough to ensure rapid convergence but small enough to insure stability. Stability of the algorithm is assured if the following condition is met (Haykin, 1996):

$$0 < \mu < \frac{2}{\lambda_{\max}} \tag{5.2.4}$$

where λ_{\max} is the largest eigenvalue of $\mathbf{R}(k)$.

The convergence time constant of the LMS algorithm depends on the eigenvalues of the correlation matrix $\mathbf{R}(k)$. Define the average eigenvalue to be:

$$\lambda_{av} = \frac{1}{M} \sum_{i=1}^M \lambda_i \tag{5.2.5}$$

Then the approximate time constant for convergence is (Haykin, 1996):

$$\tau_{mse,av} \approx \frac{1}{2\mu\lambda_{av}} \tag{5.2.6}$$

A larger learning rate parameter is desirable for shorter convergence times, but that learning parameter must not exceed the allowable limit or the LMS algorithm will not converge. In the cases shown in Figures 52 and 54, the convergence time is approximately 3 ms for $\mu = 6.31$ and 2 ms for $\mu = 10$. The sampling period is 40 ns, which implies that the convergence times in samples are approximately 50,000 samples for $\mu = 10$ and 75000 samples for $\mu = 6.31$. Using the convergence time equation above and using the number of samples for τ in this discrete-time analysis, the following average eigenvalue is inferred:

$$\lambda_{av} \approx 10^{-6} \quad (5.2.7)$$

Caution must be exercised in assuming the above average eigenvalue. Mainly, the matrix $R(k)$ is a time-dependent correlation matrix as explained below, so the stationarity assumption does not hold. Hence, the average eigenvalue shown in Eq. (5.2.7) is not correct, and a more detailed analysis is necessary. This is one of several areas where the non-stationarity of the signal will play an important role in understanding the behavior of LMS.

5.3 Non-stationarity of the input signal and convergence time

The use of time-variant autocorrelation and cross-correlation matrices R and p is not correct in general when the input signal is non-stationary. Although the additive white Gaussian noise assumption holds (Gagliardi and Karp, 1995), the desired signal $d(k)$ is not statistically stationary. In particular, the mean of the signal $d(k)$ is time dependent in the case of PPM modulation. Each PPM frame contains 512 slots, 256 of which may

contain a symbol with equal probability and 256 of which are dead time slots. Since the eigenstructure of the correlation matrix has a direct impact on LMS stability and convergence time as shown by Eqs. (5.2.4), (5.2.5), and (5.2.6), it is necessary to examine this eigenstructure in order to make meaningful statements about convergence times and learning rate parameters. The analysis in this Section focuses extensively on the eigenvalues of the time-varying correlation matrix for this reason.

According to Sayed (2003) and Honig and Messerschmitt (1984), the LMS algorithm does not require a statistically stationary input. Its time adaptive nature will allow it to converge and track even with some time variation in the input statistics. Yet it remains necessary to examine the nature of the non-stationarity of the signal under consideration, and this is considered in the analysis below. Assume that there are N FPA channels. The instantaneous ensemble autocorrelation matrix of the received N -dimensional signal vector $\mathbf{r}(k)$ modeled in Eq. (5.1.2) can be written as:

$$\mathbf{R}(k) = \mathbf{R}_s(k) + 2\sigma^2 \mathbf{I} \quad (5.3.1)$$

where

$$\mathbf{R}_s(k) = E[\mathbf{s}(k)\mathbf{s}^H(k)] \quad (5.3.2)$$

At this stage, it is necessary to analyze two distinct cases. In the first case, the following model of the signal $\mathbf{s}(k)$ will be assumed:

$$\mathbf{s}(k) = \begin{cases} \mathbf{a}d(k) & \text{signal slot} \\ \mathbf{0} & \text{noise slot} \end{cases} \quad (5.3.3)$$

In this case, since $s(k) = 0$ in the noise slots, $r(k) = n(k)$ in the noise slots since the signal contribution is zero.

In the second case, the following signal model will be assumed:

$$\mathbf{s}(k) = \begin{cases} \mathbf{a}d(k) & \text{signal slot} \\ \mathbf{b} & \text{noise slot} \end{cases} \quad (5.3.4)$$

In Eq. (5.3.4) it is assumed that a certain amount of signal energy “leaks” from the source even in the noise slots. Although this leakage energy would not be present in a realistic deep space communications system, it was present in the laboratory setup. The vector \mathbf{a} contains the complex amplitude and phase responses of the FPA detector elements to the spatial distribution of the signal in the focal plane in a PPM signal slot. The vector \mathbf{b} contains the complex amplitude responses of the FPA detector elements to the spatial distribution of the leakage in the focal plane. In the laboratory setup, the spatial distribution of the leakage was not the same as the spatial distribution of the signal, so the vectors \mathbf{a} and \mathbf{b} were not the same: these vectors were not even collinear. In Section 5.3.1 the case of no leakage, described by Eq. (5.3.3), will be covered. In Section 5.3.2 the case with leakage seen in the laboratory as described by Eq. (5.3.4) will be covered.

5.3.1 Case I: No leakage present

The matrix \mathbf{R}_s is obtained by substituting Eq. (5.3.3) into Eq. (5.3.2), resulting in:

$$\mathbf{R}_s(k) = \begin{cases} \mathbf{a}\mathbf{a}^H E\left[\left|d(k)\right|^2\right] & \text{signal slot} \\ \mathbf{0} & \text{noise slot} \end{cases} \quad (5.3.5)$$

Substitution of Eq. (5.3.5) into Eq. (5.3.1) yields:

$$\mathbf{R}(k) = \begin{cases} \mathbf{a}\mathbf{a}^H E\left[|d(k)|^2\right] + 2\sigma^2\mathbf{I} & \text{signal slot} \\ 2\sigma^2\mathbf{I} & \text{noise slot} \end{cases} \quad (5.3.6)$$

In any event, the time-dependent correlation matrix $\mathbf{R}(k)$ is expressed in the form:

$$\mathbf{R}(k) = \mathbf{R}_{s,0} |d(k)|^2 + 2\sigma^2\mathbf{I} \quad (5.3.7)$$

where:

$$\mathbf{R}_{s,0} = \mathbf{a}\mathbf{a}^H \quad (5.3.8)$$

For the moment, ignore the exact form of $\mathbf{R}_{s,0}$ specified in Eq. (5.3.8) and assume that $\mathbf{R}_{s,0}$ is a generic, non-time-varying correlation matrix. Let the set of eigenvectors of $\mathbf{R}_{s,0}$ be given by $\{\mathbf{v}_i\}$ and the set of corresponding eigenvalues be given by $\{\lambda_i\}$. Each \mathbf{v}_i is also an eigenvector of $\mathbf{R}_s = \mathbf{R}_{s,0} E\left[|d(k)|^2\right]$ with corresponding eigenvalue:

$$a_i(k) = \lambda_i E\left[|d(k)|^2\right] \quad (5.3.9)$$

In turn, each eigenvector \mathbf{v}_i of $\mathbf{R}_s E\left[|d(k)|^2\right]$ is also an eigenvector of

$\mathbf{R}(k) = \mathbf{R}_s E\left[|d(k)|^2\right] + 2\sigma^2\mathbf{I}$ with corresponding eigenvalue:

$$b_i(k) = \left(E\left[|d(k)|^2\right] + 2\sigma^2\right) \lambda_i \quad (5.3.10)$$

The set $\{\mathbf{v}_i\}$ is, therefore, a set of non-time-varying eigenvectors of the time-varying correlation matrix $\mathbf{R}(k)$ with time varying eigenvalues given by Eq. (5.3.10). Define the matrix whose columns are the constant eigenvectors:

$$\mathbf{V} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_N] \quad (5.3.11)$$

and define the diagonal matrix containing the time-varying eigenvalues:

$$\mathbf{B}(k) = \begin{bmatrix} b_1(k) & 0 & \dots & 0 \\ 0 & b_2(k) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_M(k) \end{bmatrix} \quad (5.3.12)$$

Since a correlation matrix is always Hermitian, $\mathbf{R}(k)$ can be expressed as:

$$\mathbf{R}(k) = \mathbf{V}\mathbf{B}(k)\mathbf{V}^H \quad (5.3.13)$$

Eq. (5.3.13) holds for any time-varying correlation matrix of the form specified in Eq. (5.3.7). The time dependence of the eigenvalues $b_i(k)$ will affect the learning rate parameter μ and the convergence time τ , and the effects will be determined by the eigenvalues $b_i(k)$.

For the particular $\mathbf{R}_{s,0}$ in Eq. (5.3.8), \mathbf{a} will be an eigenvector with eigenvalue $\|\mathbf{a}\|^2$ and the $(N-1)$ dimensional subspace \mathbf{a}_\perp of vectors perpendicular to \mathbf{a} will be

spanned by a $(N-1)$ dimensional basis of the remaining orthogonal eigenvectors (which are non-unique) all with corresponding eigenvalue zero. The fact that \mathbf{a} is itself an eigenvector of $\mathbf{R}(k)$, assuming Eq. (5.3.8) holds, is shown below:

$$\begin{aligned}
 \mathbf{R}(k)\mathbf{a} &= \left[\mathbf{a}\mathbf{a}^H E\left[|d(k)|^2\right] + 2\sigma^2\mathbf{I} \right] \mathbf{a} \\
 &= E\left[|d(k)|^2\right] \mathbf{a}\mathbf{a}^H \mathbf{a} + 2\sigma^2\mathbf{I}\mathbf{a} \\
 &= E\left[|d(k)|^2\right] \mathbf{a}\|\mathbf{a}\|^2 + 2\sigma^2\mathbf{a} \\
 &= \left(E\left[|d(k)|^2\right] \|\mathbf{a}\|^2 + 2\sigma^2 \right) \mathbf{a} \\
 &= \lambda_a \mathbf{a}
 \end{aligned} \tag{5.3.14}$$

where:

$$\lambda_a = \left(E\left[|d(k)|^2\right] \|\mathbf{a}\|^2 + 2\sigma^2 \right) \tag{5.3.15}$$

is the eigenvalue that corresponds to eigenvector \mathbf{a} .

Let \mathbf{a}_\perp denote the $(N-1)$ -dimensional subspaces of all vectors that are orthogonal to \mathbf{a} . Let the set $\{\mathbf{c}_i\}$ be any orthogonal basis for the subspace \mathbf{a}_\perp . Then for any vector \mathbf{c}_i from this basis:

$$\begin{aligned}
 \mathbf{R}(k)\mathbf{c}_i &= \left[\mathbf{a}\mathbf{a}^H E\left[|d(k)|^2\right] + 2\sigma^2\mathbf{I} \right] \mathbf{c}_i \\
 &= E\left[|d(k)|^2\right] \mathbf{a}\mathbf{a}^H \mathbf{c}_i + 2\sigma^2\mathbf{c}_i \\
 &= E\left[|d(k)|^2\right] \mathbf{a}(\mathbf{a}^H \mathbf{c}_i) + 2\sigma^2\mathbf{c}_i \\
 &= E\left[|d(k)|^2\right] \mathbf{a}0 + 2\sigma^2\mathbf{c}_i \\
 &= 2\sigma^2\mathbf{c}_i
 \end{aligned} \tag{5.3.16}$$

By Eq. (5.3.16) it is now clear that the vectors \mathbf{c}_i must also be eigenvectors of $\mathbf{R}(k)$. No matter how the set $\{\mathbf{c}_i\}$ of orthogonal vectors spanning \mathbf{a}_\perp is chosen, the set of eigenvalues of $\mathbf{R}(k)$ will be given by:

$$b_i(k) = \begin{cases} 2\sigma^2 + \|\mathbf{a}\|^2 E\left[\left|d(k)\right|^2\right] & i = 1 \\ 2\sigma^2 & i \geq 2 \end{cases} \quad (5.3.17)$$

Only the largest eigenvalue of $\mathbf{R}(k)$ is a function of time, but it is the largest eigenvalue that will also determine the upper limit on the learning rate parameter μ as well as the convergence time constant. Since Eq. (5.2.4) yields an upper bound on the learning rate parameter, and since the location of the signal pulse is not known until demodulation is performed, it is necessary to bound the learning rate parameter using the largest possible value of λ_{\max} , which is:

$$\lambda_{\max} = 2\sigma^2 + \|\mathbf{a}\|^2 d_{\max}^2 \quad (5.3.18)$$

This yields the following upper bound on the learning rate parameter:

$$\mu < \frac{2}{2\sigma^2 + \|\mathbf{a}\|^2 d_{\max}^2} \quad (5.3.19)$$

Furthermore, the average eigenvalue used in Eq. (5.2.6) depends on the average eigenvalue from Eq. (5.2.5). A substitution of the eigenvalues from Eq. (5.3.17) into Eq. (5.2.5) yields:

$$\lambda_{\text{avg}} = 2\sigma^2 + \frac{\|\mathbf{a}\|^2 d_{\max}^2}{N} \quad (5.3.20)$$

Replacing μ in Eq. (5.2.6) with the upper bound from Eq. (5.3.19) and plugging in Eq. (5.3.20) yields:

$$\tau \approx \frac{2\sigma^2 + \|\mathbf{a}\|^2 d_{\max}^2}{4 \left(2\sigma^2 + \frac{\|\mathbf{a}\|^2 d_{\max}^2}{N} \right)} \quad (5.3.21)$$

In the limit of high SNR when $\|\mathbf{a}\|^2 d_{\max}^2 \gg N\sigma^2$, Eq. (5.3.21) can be further approximated by:

$$\begin{aligned} \tau &\approx \frac{2\sigma^2 + \|\mathbf{a}\|^2 d_{\max}^2}{4 \left(2\sigma^2 + \frac{\|\mathbf{a}\|^2 d_{\max}^2}{N} \right)} \\ &\approx \frac{\|\mathbf{a}\|^2 d_{\max}^2}{\left(\frac{4}{N} \right) \|\mathbf{a}\|^2 d_{\max}^2} \\ &= \frac{N}{4} \end{aligned} \quad (5.3.22)$$

As stated previously, N is the number of FPA channels. In principle, for a 16 channel system at sufficiently high SNR, convergence times can be as rapid as only 4 samples. However, the convergence time constant shown in Eq. (5.3.22) is based on Eq. (5.2.6).

Eq. (5.3.18) is an upper bound on the maximum eigenvalue, and Eq. (5.3.20) is an upper bound on the average eigenvalue. These upper bounds hold within the signal slots. Within the noise slots, the maximum and average eigenvalues are both equal to $2\sigma^2$,

resulting in even larger maximum allowable values for μ (i.e., $\mu < \frac{1}{2\sigma^2}$) that are very large for small noise variance.

Since the goal of a PPM communications system is to determine the location of the signal slot, the signal slot's location is not known *a priori*. Hence, the upper bound for the learning rate parameter μ is necessarily given by the largest possible eigenvalue as shown in Eq. (5.3.19). For the data set shown from Figure 54 through Figure 57, the magnitude of the term $\|\mathbf{a}\|^2 d_{\max}^2$ is $1.3064 \cdot 10^{-3}$. The maximum learning rate parameter that may be used throughout the entire data sequence in both signal and noise slots is thus given by the upper bound from the signal slots from Eq. (5.3.19), which yields:

$$\mu_{\max} \approx 1.53 \cdot 10^3 \quad (5.3.23)$$

Therefore it is no surprise that stable convergence is achieved in the case when $\mu = 1000$ in Figure 56 but that instability occurs when $\mu = 10,000$ in Figure 57 in light of the upper bound in Eq. (5.3.23). Furthermore, the nearly instantaneous convergence of the system is also not a surprise. Since $N = 4$ in the figures above, Eq. (5.3.22) would suggest that within a signal slot the system should converge in only about 1 sample of training. This is clearly in evidence in Figure 51 because the output pulse in red is very nearly equal to the target pulse in blue even for the very first PPM pulse.

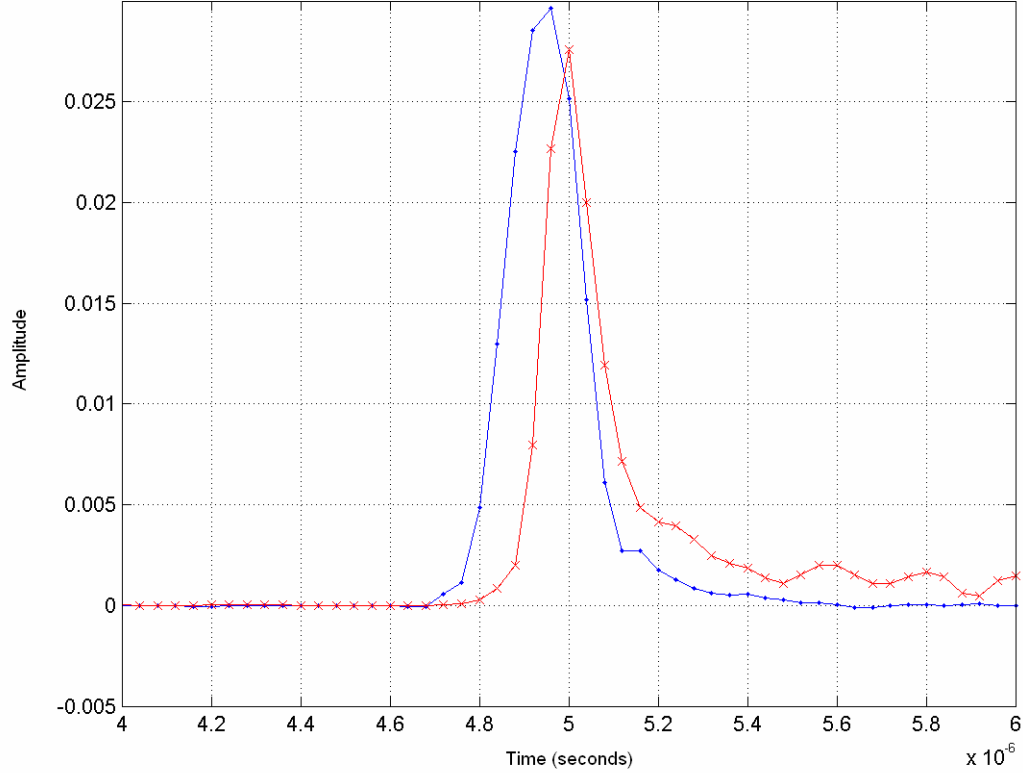


Figure 51. Magnitude of the reference signal (blue) and the LMS output (red). The LMS output shows an apparent delay because it takes one or two samples for the LMS algorithm to converge, as predicted by Eq. (5.3.22) for $N = 4$.

For smaller μ , the LMS algorithm's convergence performance will be slower. If one assumes $\mu = 10$ for example, then Eq. (5.2.6) along with $\|\mathbf{a}\|^2 d_{\max}^2 = 1.3064 \cdot 10^{-3}$ would yield an approximate convergence time of 38 samples. However, there are only about five or six large samples per pulse in this case, so it would take at least 5 to 8 PPM symbols for convergence to occur. Indeed, this appears to be roughly consistent with the fact that convergence seems to occur primarily within the first 10 pulses in Figure 54

where the value $\mu = 10$ is used. Since the analysis here is very approximate due to the fact that this is a shaped pulse rather than a rectangular pulse, only very rough numbers can be obtained for convergence time. Nevertheless, the fact that convergence times in the experimental data series roughly agree with the expected times found here should be encouraging.

Note that convergence with $\mu = 10$ is notably slower due to the presence of non-stationarity. In Chapter 4, one notes that LMS convergence times for M -ary PPM were approximately M times slower than they were for constant signals. This is reasonable in light of the fact that the LMS time constant of convergence in Eq. (5.2.6) is inversely proportional to the average eigenvalue. In the noise slots, the average eigenvalue is only $2\sigma^2$ while in the signal slots the average eigenvalue is $2\sigma^2 + \frac{\|\mathbf{a}\|^2 d_{\max}^2}{N}$, which is significantly larger for high SNR. Indeed, if $\|\mathbf{a}\|^2 d_{\max}^2 \gg 2\sigma^2 N$ then one can state:

$$\begin{aligned}\tau_{sig} &= \frac{1}{2\mu \left(\frac{\|\mathbf{a}\|^2 d_{\max}^2}{N} + 2\sigma^2 \right)} \\ \tau_{noise} &= \frac{1}{4\mu\sigma^2} \\ \left(\|\mathbf{a}\|^2 d_{\max}^2 \gg 2\sigma^2 N \right) &\Rightarrow \left(\tau_{sig} \ll \tau_{noise} \right)\end{aligned}\tag{5.3.24}$$

In the high SNR case, convergence is largely dominated by the signal slots. Hence, the weights undergo useful convergence for a fraction of the total running time equal to $\frac{1}{M}$, and this is entirely consistent with the above analysis.

The type of non-stationarity encountered in the case of PPM modulation will, therefore, lead to a factor of M decrease in convergence times. This is supported by theoretical analysis, by the simulation results of Chapter 4, and by some observations on convergence times with real-world PPM data.

5.3.2 Case II: Leakage present

The analysis with the presence of leakage necessarily changes because the vectors \mathbf{a} and \mathbf{b} are not parallel due to the fact that the leakage has a different complex spatial distribution in the focal plane than the desired signal. Experimental evidence for this is provided in the Appendix. In particular, it is no longer possible to claim that the eigenvectors of the correlation matrix are constant as a function of time because this is no longer true. Following an analysis similar to that in Section 5.3.1, it can be stated that in the noise slots the largest eigenvalue is given by:

$$\lambda_b = 2\sigma^2 + \|\mathbf{b}\|^2 \quad (5.3.25)$$

with \mathbf{b} itself as the corresponding eigenvector. The subspace of vectors orthogonal to vector \mathbf{b} , which is denoted here as \mathbf{b}_\perp , will not be the same as the subspace \mathbf{a}_\perp . However, the following statements can still be made. First, the leakage vector \mathbf{b} is of much lower magnitude than the signal vector \mathbf{a} . This implies:

$$\|\mathbf{a}\|^2 d_{\max}^2 \gg \|\mathbf{b}\|^2 \quad (5.3.26)$$

Since the highest eigenvalue in a signal slot given by Eq. (5.3.18) is significantly greater than the maximum eigenvalue in a noise slot given by Eq. (5.3.25), the upper bound on the parameter μ in Eq. (5.3.19) does not change at all because it is necessary to

choose μ to insure stability in both signal and noise slots. Hence, Eq. (5.3.23) remains entirely valid even in the laboratory case with leakage present. Furthermore, since leakage does not change convergence characteristics within a signal slot, where no leakage is present, the convergence time analysis from case I remains applicable as well. Hence, the eigenstructure of the time-varying correlation matrix within a signal slot is still the dominant factor in determining convergence behavior for LMS with a reliable training sequence. Eq. (5.3.26) would imply that LMS convergence is again dominated, to a large extent, by the signal slots. Hence, even with leakage present one expects to observe a factor of M slowdown in convergence for an M -ary PPM system when compared with a system with a constant signal.

5.4 More on the effects of leakage on the eigenstructure

Although leakage did not significantly impact the convergence time analysis, it still has important eigenstructure effects as explained in Section 5.3. In particular, the leakage causes changes in the eigenvectors of the correlation matrix as a function of time. Eq. (5.2.2) is repeated again below:

$$\begin{aligned}
 \nabla_{\mathbf{w}(k)} \text{MSE} &= -2\mathbf{p} + 2\mathbf{R}\mathbf{w}(k) \\
 &= -2E[\mathbf{r}(k)d^*(k)] + 2E[\mathbf{r}(k)\mathbf{r}^H(k)]\mathbf{w}(k) \\
 &\approx -2\mathbf{r}(k)(d(k) - \mathbf{w}^H(k)\mathbf{r}(k))^* \\
 &= -2\mathbf{r}(k)e^*(k)
 \end{aligned} \tag{5.4.1}$$

There are two important sets of eigenvectors that must now be considered. First, let $\{\mathbf{v}_{i,\text{sig}}\}$ be the set of eigenvectors of the signal slot correlation matrix given by:

$$\mathbf{R}_{sig}(k) = E \left[\mathbf{a} \mathbf{a}^H E \left[|d(k)|^2 \right] + 2\sigma^2 \mathbf{I} \right] \quad (5.4.2)$$

And let $\{\mathbf{v}_{i,leak}\}$ be the set of eigenvectors of the leakage correlation matrix given by:

$$\mathbf{R}_{leak}(k) = \mathbf{b} \mathbf{b}^H + 2\sigma^2 \mathbf{I} \quad (5.4.3)$$

Let the first eigenvector of each set be the one corresponding to the largest eigenvalue. This yields:

$$\begin{aligned} \mathbf{v}_{1,sig} &= \mathbf{a} \\ \mathbf{v}_{1,leak} &= \mathbf{b} \end{aligned} \quad (5.4.4)$$

At this point, it helps to note that $d(k) = 0$ in noise slots and that:

$$\mathbf{r}(k) = \begin{cases} \mathbf{a}d(k) + \mathbf{n}(k) & \text{signal slot} \\ \mathbf{b} + \mathbf{n}(k) & \text{noise slot} \end{cases} \quad (5.4.5)$$

This gives:

$$\begin{aligned} \mathbf{p}(k) &= E \left[\mathbf{r}(k) d^*(k) \right] \\ &= \begin{cases} E \left[(\mathbf{a}d(k) + \mathbf{n}(k)) d^*(k) \right] & \text{signal slot} \\ E \left[(\mathbf{b} + \mathbf{n}(k)) 0 \right] & \text{noise slot} \end{cases} \\ &= \begin{cases} \mathbf{a} E \left[|d(k)|^2 \right] & \text{signal slot} \\ \mathbf{0} & \text{noise slot} \end{cases} \end{aligned} \quad (5.4.6)$$

The reason $\mathbf{p}(k) = 0$ in a noise slot stems from two facts. First, $d(k) = 0$ whenever we are in a noise slot, so within a noise slot $E[d(k)] = 0$. Second, it is assumed that the AWGN is zero mean and entirely uncorrelated with the desired signal whether in a signal slot or in a noise slot, which implies $E[\mathbf{n}(k) d^*(k)] = \mathbf{0}$. The optimal weights for the

noise slots are zeros. Since the locations of the signal slots and of the noise slots are not known *a priori*, it is not possible to set the weights to zero for each noise slot. Here, it will be assumed that the optimal weights for the signal slot are available. These weights are known to satisfy:

$$\mathbf{R}_{sig}(k) \mathbf{w}_{opt}(k) = \mathbf{p}_{sig}(k) \quad (5.4.7)$$

If

$$\mathbf{w} = \alpha \mathbf{a} \quad (5.4.8)$$

then the previous eigenvector and eigenvalue analysis of the correlation matrix yields:

$$\begin{aligned} \mathbf{R}_{sig}(k) \mathbf{w} &= \alpha \mathbf{R}_{sig}(k) \mathbf{a} \\ &= \alpha \left(2\sigma^2 + \|\mathbf{a}\|^2 E\left[|d(k)|^2\right] \right) \mathbf{a} \end{aligned} \quad (5.4.9)$$

Setting

$$\alpha = \frac{E\left[|d(k)|^2\right]}{\left(2\sigma^2 + \|\mathbf{a}\|^2 E\left[|d(k)|^2\right] \right)} \quad (5.4.10)$$

Yields:

$$\begin{aligned} \mathbf{R}_{sig}(k) \mathbf{w} &= \alpha \mathbf{R}_{sig}(k) \mathbf{a} \\ &= \alpha \left(2\sigma^2 + \|\mathbf{a}\|^2 E\left[|d(k)|^2\right] \right) \mathbf{a} \\ &= \mathbf{a} E\left[|d(k)|^2\right] \\ &= \mathbf{p}(k) \end{aligned} \quad (5.4.11)$$

which is the Wiener solution. Since a constant scaling factor in the weights is unimportant, the best weight vector within the signal slot is really the largest eigenvector

a. Setting $\mathbf{w} = \mathbf{a}$ therefore yields the Wiener solution to within a scalar factor. It must be emphasized at this point that the claim that the largest eigenvector \mathbf{a} contains the optimal weights is dependent on Eq. (5.4.2) (or, equivalently, Eq. (5.3.8)) holding true.

Assume now that the weights are kept fixed at $\mathbf{w} = \mathbf{a}$, the optimal weight vector for the signal slots. This will have an effect on the gradient of the mean squared error in the noise slots. In particular, substituting into Eq. (5.2.1) but with the correlation matrix $\mathbf{R}(k)$ for the noise and leakage used in place of the correlation matrix for the desired signal and with $\mathbf{p}(k) = 0$ will yield the following gradient within the noise slots:

$$\begin{aligned}\nabla_{\mathbf{w}} \text{MSE}(\mathbf{k}) &= -2\mathbf{p}(k) + 2\mathbf{R}_{\text{noise}}(k)\mathbf{w}(k) \\ &= -\mathbf{0} + 2(\mathbf{b}\mathbf{b}^H + 2\sigma^2\mathbf{I})\mathbf{a} \\ &= 4\sigma^2\mathbf{a} + 2(\mathbf{b}^H\mathbf{a})\mathbf{b}\end{aligned}\tag{5.4.12}$$

Eq. (5.4.12) indicates that the gradient will be a linear combination of the vectors \mathbf{a} and \mathbf{b} . This has the following implication: the updates of the LMS weights that are carried out within the leakage slots will, in general, not be parallel to the vector \mathbf{a} . Hence, even in those cases where the system has converged very close to the optimum weights within the signal slot, there will be a tendency to drift *away* from the optimal weights during the updates in the noise slots. Depending on the strength of the leakage as given by the magnitude of the inner product $\mathbf{b}^H\mathbf{a}$, this could result in significant performance degradation. The effect has the potential to be very pronounced in 256 PPM since 255 out of 256 slots will have noise and leakage without the desired signal. If LMS is continuously applied to a 256 PPM with 256 dead time slots in addition to 256 potentially valid signal slots, the problem is compounded because the system spends $\frac{511}{512}$

of its time performing a gradient update that takes it away from the SNR maximizing weights within the signal slot.

Note that if $\mathbf{b} = 0$ in Eq. (5.4.12) above, the no leakage case is obtained again and the weight update is parallel to the optimum weight vector \mathbf{a} . Hence, the absence of leakage would help to insure that this tendency to drift away from the optimal weights would be largely avoided, except for the natural wandering of the weights near the optimal set (Sayed, 2003). Although BER results still indicate good LMS combining performance even with leakage presence, it will be seen that the leakage term negatively impacts a constant target version of the LMS algorithm considered next.

5.5 LMS with a constant target

A constant target variant of the LMS algorithm is described in Compton, (1988). If the desired signal $d(k)$ does not have a zero mean, then it possesses a correlation with a constant target. It has been shown that it is possible to replace the true desired signal $d(k)$ with a constant (Compton, 1988). Two cases will be analyzed in this section: case I without leakage present and case II with leakage present.

If no signal leakage is present, then the LMS update Eq. with a constant target can be written as:

$$\begin{aligned}
E[\mathbf{w}(k+1)] &= E[\mathbf{w}(k) + \mu e^*(k) \mathbf{r}(k)] \\
&= E[\mathbf{w}(k) + \mu \mathbf{r}(k) [C - y(k)]^*] \\
&= \mathbf{w}(k) + E\left[\mu [\mathbf{a}d(k) + \mathbf{n}(k)] [C - \mathbf{w}^H(k) \mathbf{r}(k)]^*\right] \\
&= \mathbf{w}(k) + \mu E\left[[\mathbf{a}d(k) + \mathbf{n}(k)] \left[C - [\mathbf{a}d(k) + \mathbf{n}(k)]^H \mathbf{w}(k)\right]\right] \\
&= \mathbf{w}(k) + \mu CE[\mathbf{r}(k)] - \mu \mathbf{R}_{sig}(k) \mathbf{w}(k) \\
&= \mathbf{w}(k) + \mu \mathbf{a}CE[d(k)] - \mu \mathbf{R}_{sig}(k) \mathbf{w}(k)
\end{aligned} \tag{5.5.1}$$

From Eq. (5.5.1) it can be shown that the gradient can be set to zero by setting the weights equal to an appropriately scaled multiple of the eigenvector \mathbf{a} . In other words, for sufficiently small μ , Eq. (5.5.1) can converge in the mean to a scalar multiple of the optimal weight vector \mathbf{a} . This is an important result: if a reliable reference signal is not available then a constant reference signal may be used to obtain convergence of the weights. The above analysis will work when no leakage is present. If the weights are already a scalar multiple of \mathbf{a} , then the above Eq. will tend to keep them there, although some noise will be present. In the noise slots, $E[d(k)] = 0$ and $\mathbf{R}_{noise}(k) \mathbf{a} = 2\sigma^2 \mathbf{a}$, so there will not be a significant tendency at all to drift away from a scalar multiple of \mathbf{a} provided that μ is kept sufficiently small. Only a certain amount of random weight noise will be present, as in “traditional” LMS.

The leakage case, however, is significantly worse. The update Eq. in the noise slots is given by:

$$\begin{aligned}
E[\mathbf{w}(k+1)] &= E[\mathbf{w}(k) + \mu e^*(k) \mathbf{r}(k)] \\
&= E[\mathbf{w}(k) + \mu \mathbf{r}(k) [C - y(k)]^*] \\
&= \mathbf{w}(k) + E[\mu [\mathbf{b} + \mathbf{n}(k)] [C - \mathbf{w}^H(k) \mathbf{r}(k)]^*] \\
&= \mathbf{w}(k) + \mu E[\mu [\mathbf{b} + \mathbf{n}(k)] [C - [\mathbf{b} + \mathbf{n}(k)]^H \mathbf{w}(k)]] \\
&= \mathbf{w}(k) + \mu CE[\mathbf{r}(k)] - \mu \mathbf{R}_{leak}(k) \mathbf{w}(k) \\
&= \mathbf{w}(k) + \mu C\mathbf{b} - \mu \mathbf{R}_{leak}(k) \mathbf{w}(k)
\end{aligned} \tag{5.5.2}$$

Repeated iteration of Eq. (5.5.2) will cause the weights to drift away from the direction of \mathbf{a} and toward the direction of \mathbf{b} , which is not good. This problem also exists in LMS with a standard training sequence. At this point, it is useful to evaluate the gradient term with $\mathbf{w} = \mathbf{a}$. Doing so yields:

$$\mu(C\mathbf{b} - \mathbf{b}\mathbf{b}^H\mathbf{a}) = \mu(C - \mathbf{b}^H\mathbf{a})\mathbf{b} \tag{5.5.3}$$

It is useful to compare Eq. (5.5.3) to Eq. (5.4.12). The term $(\mathbf{b}^H\mathbf{a})\mathbf{b}$ is the same in both Eqs., but the term $2\sigma^2\mathbf{a}$ in Eq. (5.4.12) is replaced by the term $C\mathbf{b}$ in Eq. (5.5.3). In other words, *the constant target LMS algorithm will exhibit a greater tendency to drift toward vector \mathbf{b} than the traditional LMS algorithm will.* This is an important point. The blind, constant target based updates will tend to reinforce the term $E[\mathbf{r}(k)]$ in Eq. (5.5.2) blindly, and since $E[\mathbf{r}(k)] = \mathbf{b}$ in the noise slots when leakage is present, this will result in notably degraded performance in the leakage case. Hence, constant target

LMS works well in simulations in which leakage is not present but will often fail to work well in real-world cases where leakage is present.

5.6 LMS convergence with real world data

The difference between LMS with a known training sequence and decision feedback LMS is that the desired signal is computed based on receiver symbol decisions. For the experiment under study, the receiver had perfect knowledge of the first 10 PPM symbols that were used as a training sequence. This training sequence was used during the acquisition phase, and once this stage is completed it is assumed that PPM symbol decisions are reliable. The remaining PPM symbols of the target signal are obtained using decision feedback. Symbol decisions from the demodulator are used to compute the target sequence in a symbol-by-symbol fashion via remodulation. The new modulated sequence, assuming correct symbol decisions, is the desired signal $d(k)$. If this signal is used as the target sequence in the LMS algorithm, then provided that no symbol decision errors occur, there will be no difference between decision feedback LMS and LMS with a completely known training sequence. When decision errors occur, however, the behavior of decision feedback LMS will not be identical to that of LMS with a perfect target sequence: errors in the target sequence will negatively impact convergence performance. Nevertheless, since communications is the primary objective, target sequences are unknown in general. Hence, decision feedback is a practical necessity that is extensively used in real-world communication systems (Proakis, 2000).

If very large learning rate parameters are used in a decision feedback LMS system, a single decision error is sufficient to cause instability in the weights and in the outputs, particularly in the case of PPM. If the learning rate parameter is kept sufficiently small, however, the weights will change relatively little even when a small number of errors are made in the decision feedback process. Hence, one expects decision feedback based LMS systems to be less stable with high learning rate parameters than LMS systems with perfectly known target sequences.

An example of the differences in stability between decision feedback LMS and LMS with a perfectly known target sequence is shown here. The behavior of decision feedback LMS with learning rate parameters $\mu = 6.31$ and $\mu = 7.94$ is illustrated in Figures 52 and 53, respectively.

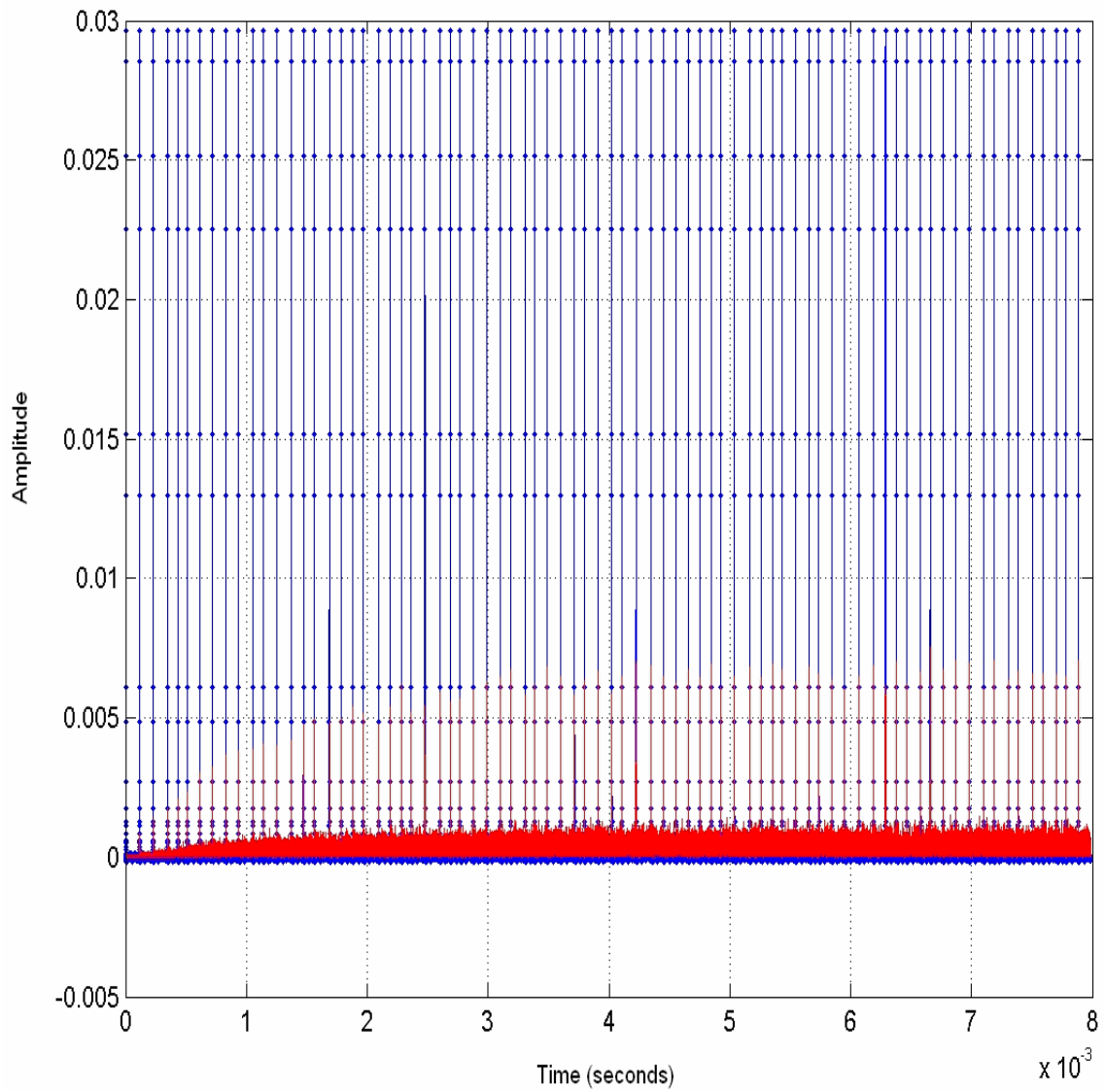


Figure 52. LMS convergence with decision feedback and with $\mu = 6.31$.

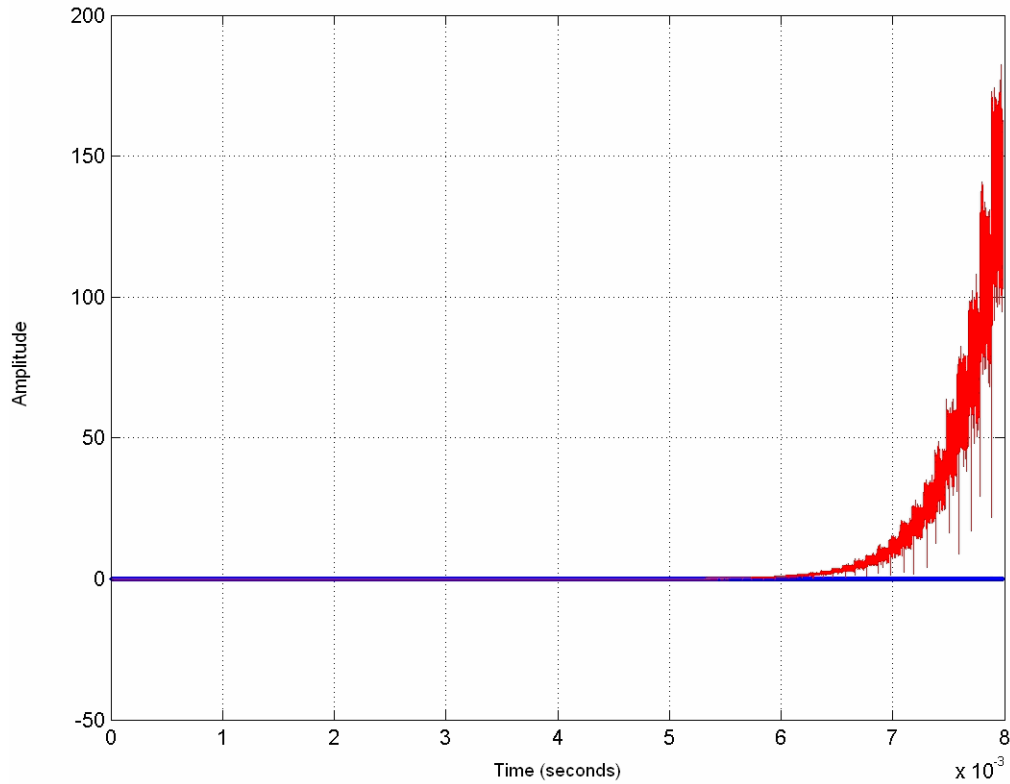


Figure 53. LMS instability with $\mu = 7.94$ and decision feedback.

When a perfectly known training sequence is used throughout, LMS exhibits significantly different behavior. This behavior is illustrated for $\mu = 10$ in Figure 54. Note that LMS remains entirely stable due to the use of a reliable training sequence. Indeed, LMS with an ideal training sequence will remain stable even at much higher learning rate parameters as illustrate for $\mu = 100$ in Figure 55. However, LMS will eventually become unstable even with an ideal training sequence as illustrated for $\mu = 10,000$ in Figure 57.

The theoretical treatments of the LMS algorithm in the literature normally assume an ideal target sequence, and even with such a sequence there is an upper bound on allowable LMS parameters as explained in previous sections.

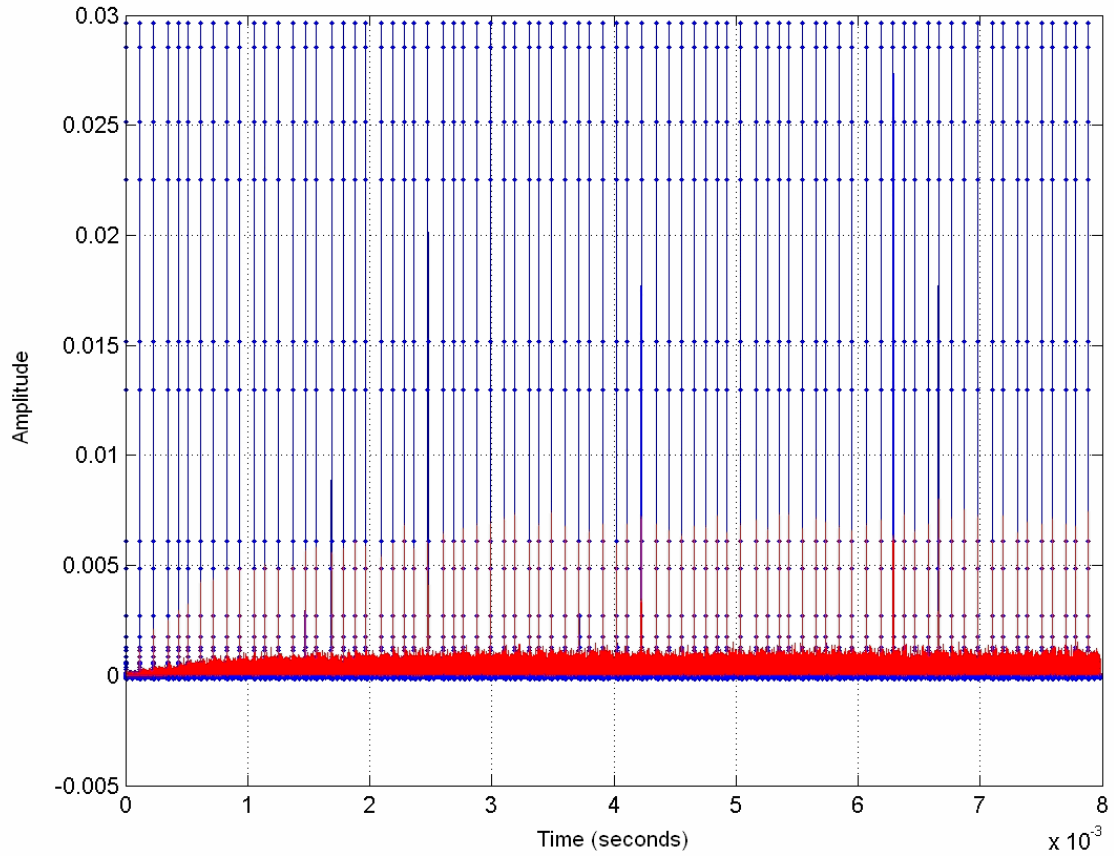


Figure 54. The behavior of LMS with $\mu = 10$ but with a perfect training sequence.

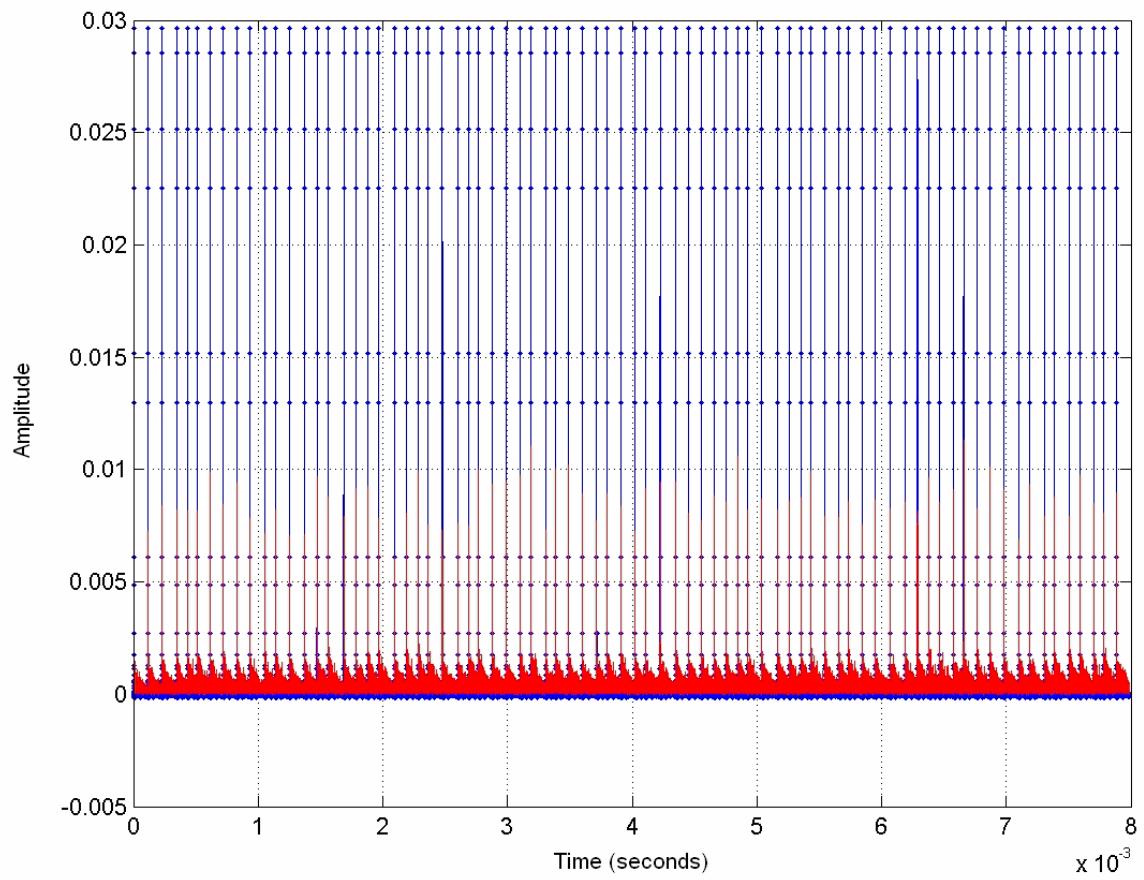


Figure 55. LMS Convergence with $\mu = 100$ but with a perfect training sequence.

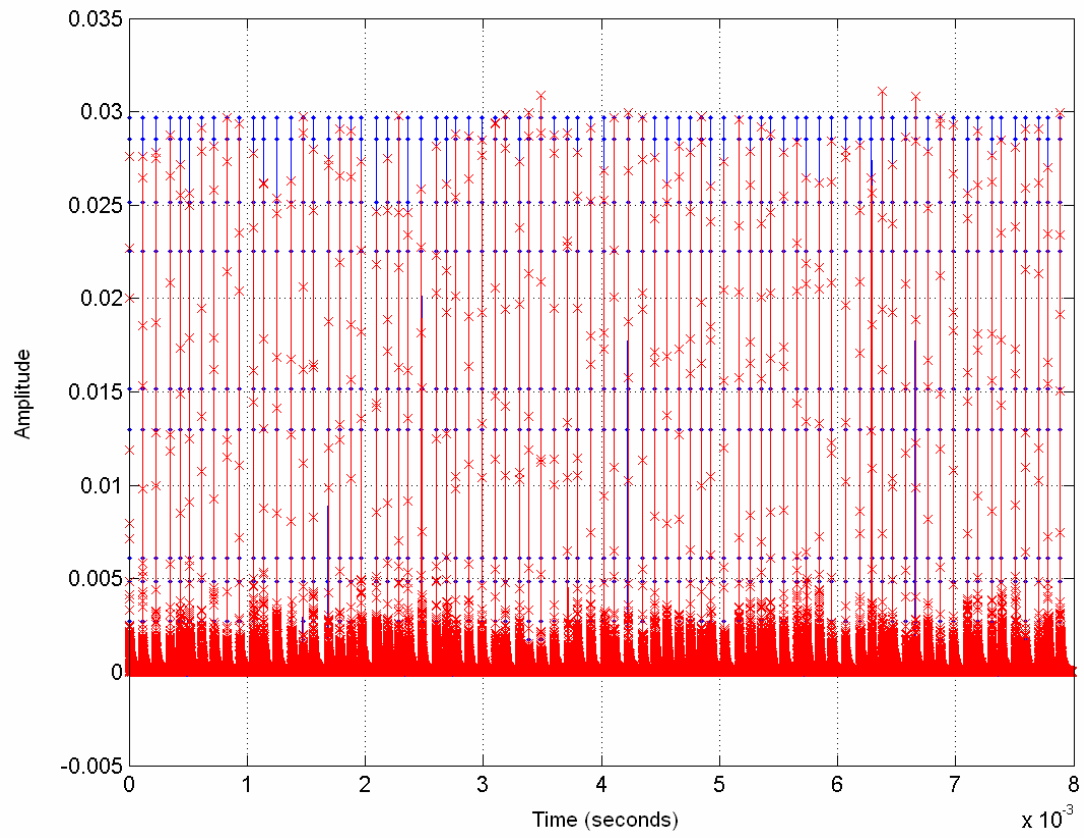


Figure 56. Convergence with $\mu = 1000$ and with a perfect target sequence.

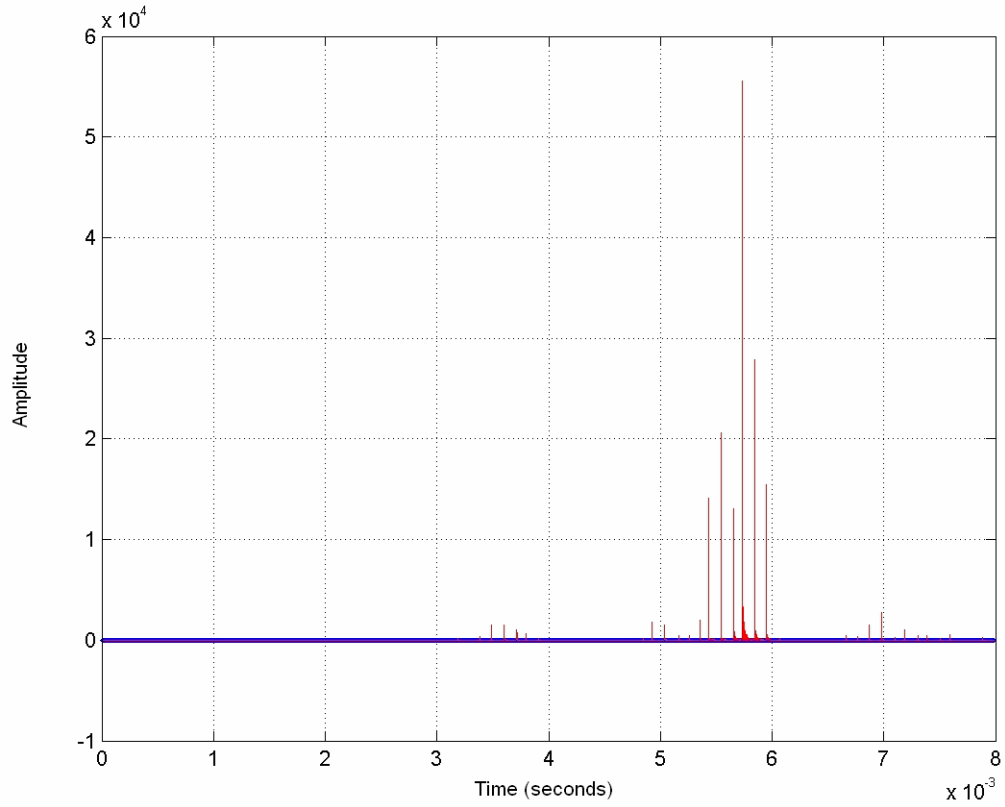


Figure 57. LMS loss of stability at $\mu = 10,000$. Perfect training sequences are of limited value here.

The convergence time of the LMS algorithm is a key issue in determining the learning rate parameter. As shown in Figure 52, the LMS algorithm was able to acquire within 3 ms with 256 PPM modulation. This represents the fastest realistic convergence time under this experimental setup for the case of decision feedback, as larger μ parameters lead to a loss of stability in the decision feedback case.

Although amplitude convergence of the LMS algorithm is important, so too is phase convergence. Since the desired signal $d(k)$ is, in this case, a real-valued PPM

signal, the output $y(k) = \mathbf{w}(k)^H \mathbf{r}(k)$ should ideally be real-valued and close to $d(k)$ as well. It is known that,

$$\begin{aligned} E[\mathbf{r}(k)] &= E[\mathbf{a}d(k) + \mathbf{n}(k)] \\ &= \mathbf{a}E[d(k)] + \mathbf{0} \\ &= \mathbf{a}E[d(k)] \end{aligned} \tag{5.6.1}$$

The phases of the above expectation are simply the phases of the terms in the vector \mathbf{a} . It is argued, but not proven here, that the phases of the output process $\mathbf{r}(k)$ are therefore simply noise versions of the phases of the optimal weights. Put another way, the phases of the optimal weights should be equal to the expected phases of the input channels to which they correspond. The optimal Wiener solution has this property and, since the LMS algorithm converges to the Wiener solution, the weights computed by the LMS algorithm should also possess this property.

There is experimental evidence that the weights generated by the LMS algorithm do track the phases well. Following are four figures that constitute LMS phase performance snapshots with $\mu = 100$. In each of the following figures, the random (due to noise) channel phase is plotted using blue dots, and the phase of the weights computed by the LMS algorithm is plotted using a red line. Each plot corresponds to a single channel.

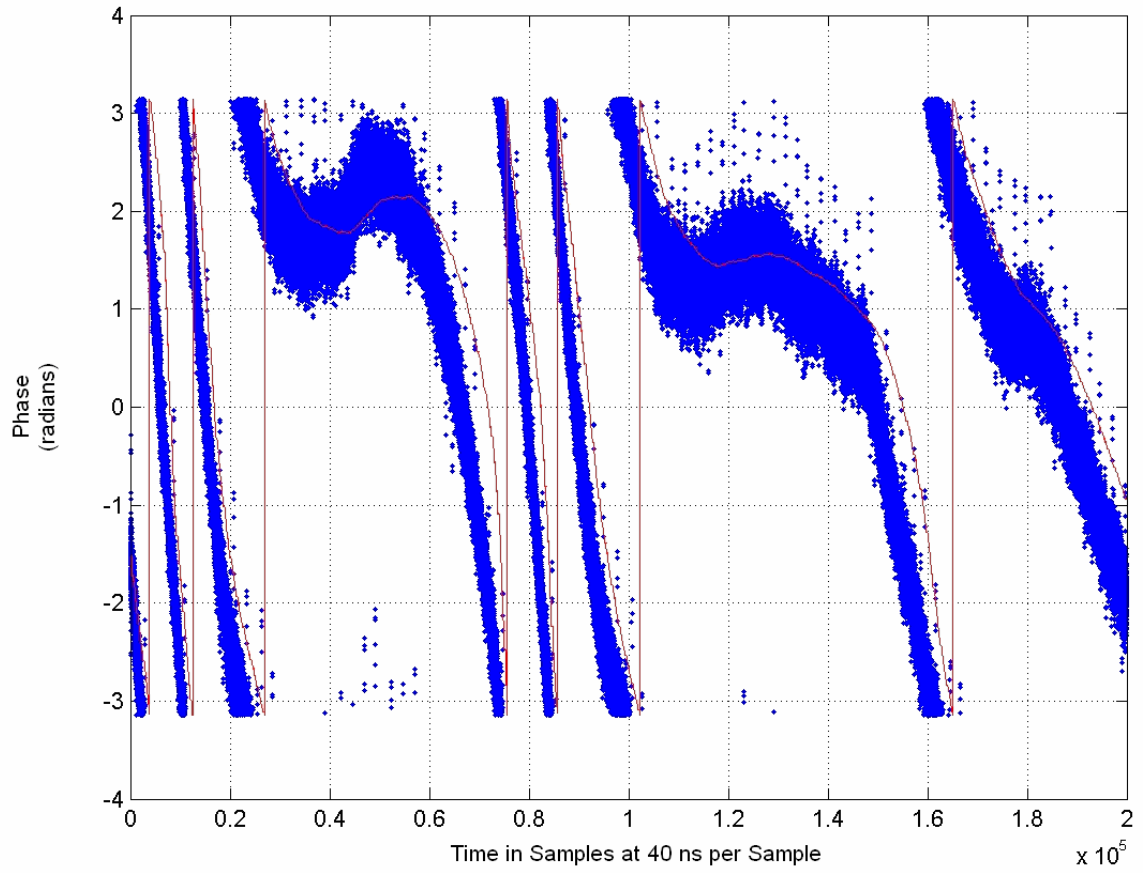


Figure 58. Tracking of the input phase of channel 1 by the first weight.

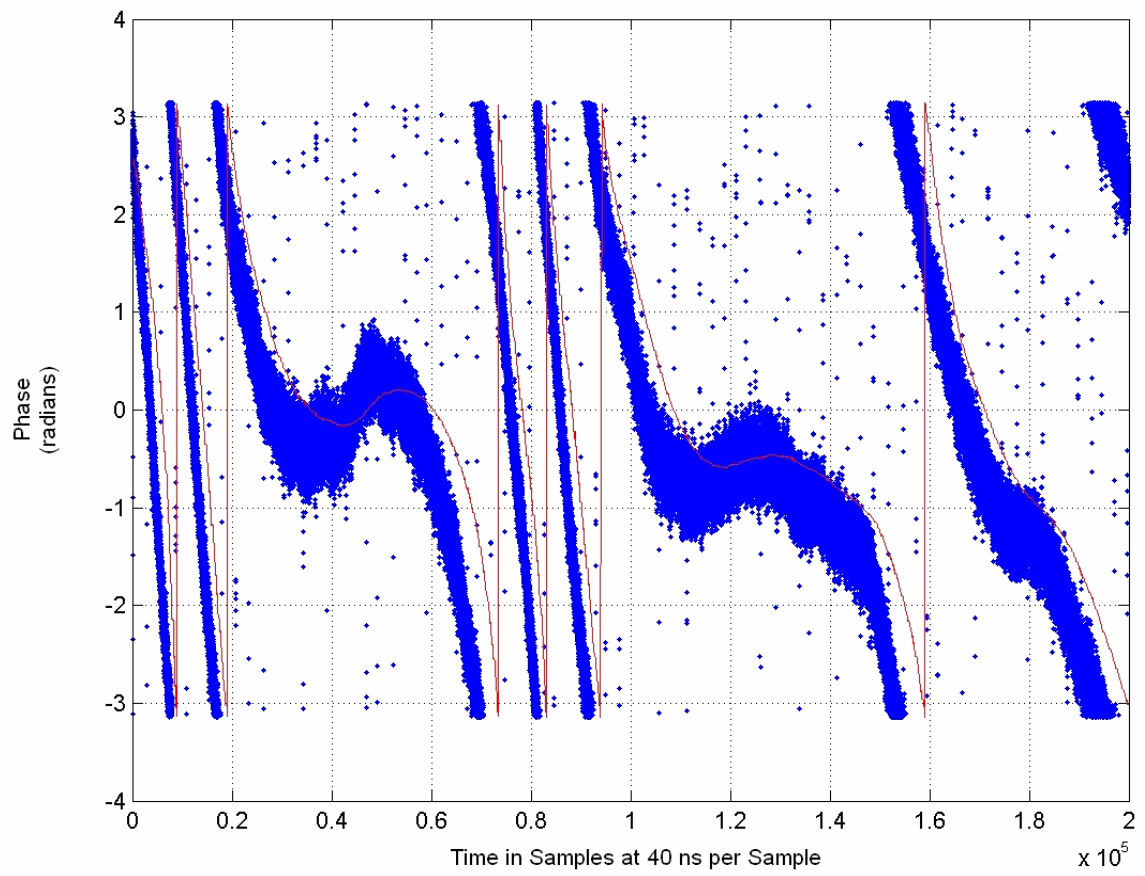


Figure 59. Tracking of the input phase of channel 2.

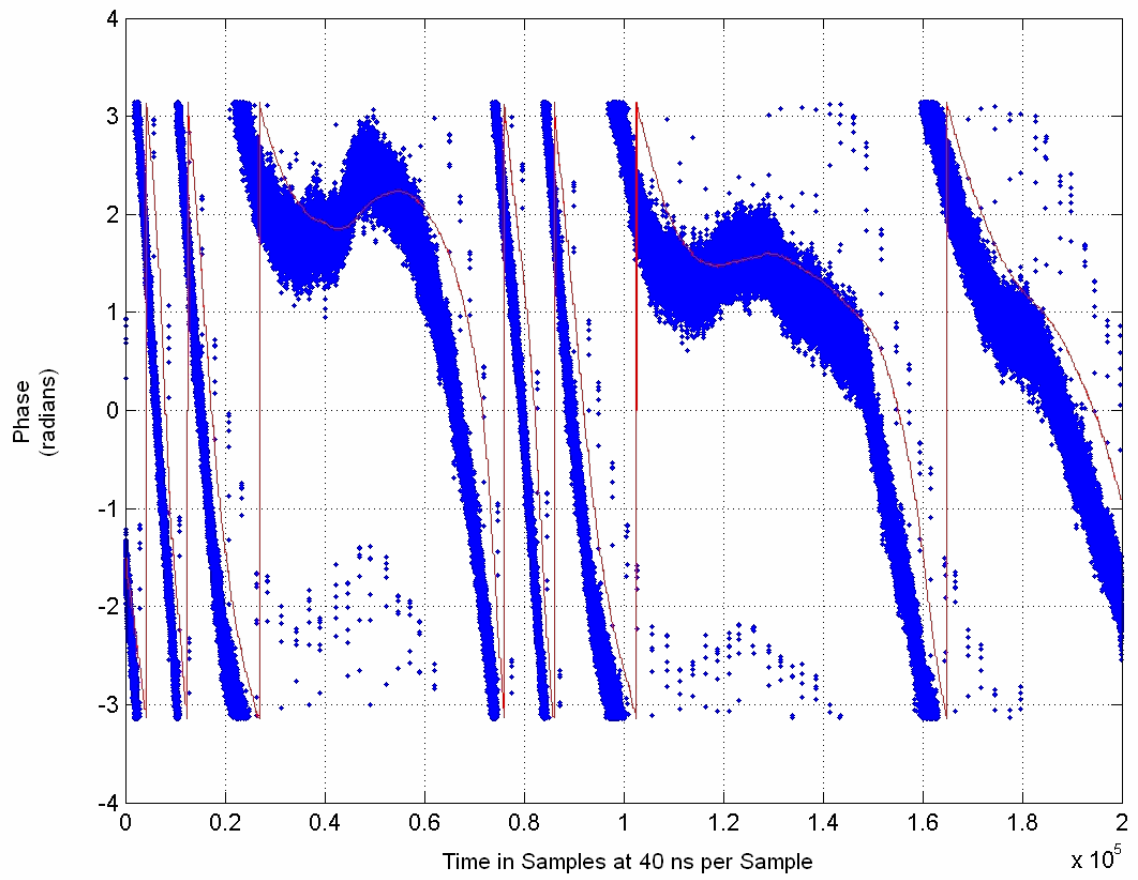


Figure 60. Tracking of the input phase of channel 3.

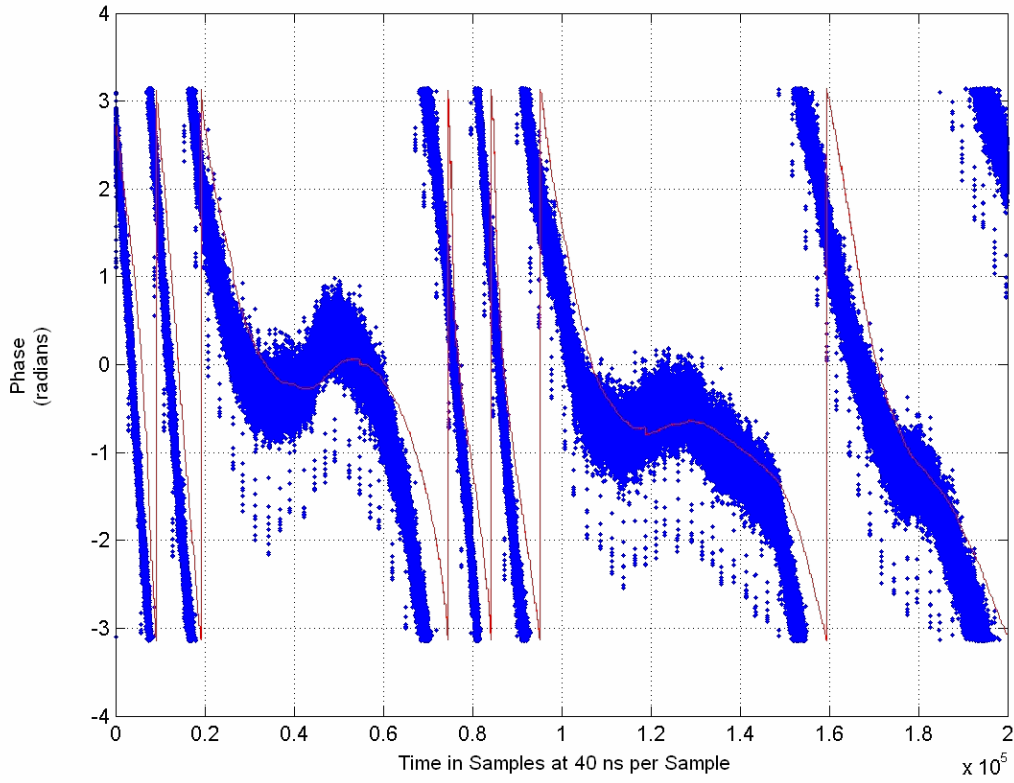


Figure 61. Tracking of the input phase of channel 4.

Since the four plots above correspond to four physical FPA channels, there is definite phase coherence among them even though their absolute phases differ. In each of the four cases, the LMS algorithm's weights tracking the phase, and the evidence for this comes from the red line denoting weight phase tracking the blue regions indicating channel input phase.

Chapter 6: Investigation of Performance of the Coherent Optical Receiver Experiment (CORE): Comparison of Experimental Results with Theory

6.1 Overview and methodology

The ability to achieve good bit error rate (BER) performance for weak signals received in the presence of atmospheric turbulence and the ability to coherently combine multiple channels with the LMS algorithm were both verified experimentally. All of the results presented here are for uncoded bit error rates. The available equipment did not enable gathering of coded BER data. Furthermore, limitations on the quantity of data that could be gathered would have made it impractical to gather the sheer volumes of experimental data required to generate reliable coded BER curves, and these limitations are explained later. In this section, both the experimental methodology and the manner in which results were generated are covered.

Two forms of PPM modulation were used to obtain experimental BER data.

1. Fixed 32 PPM. In this PPM version, a constant symbol of 0 (with possible symbols 0 through 31) was sent in each PPM frame. Each slot consisted of eight samples, and there were 32 slots for a total of 256 samples per PPM frame. Each data set contained 779 PPM symbols.

2. Random 256 PPM. In this case, a random set of integers from 0 through 255 were modulated on the PPM frames. Each frame consisted of 512 slots, with 256 potential signal slots and 256 dead time slots. Each slot contained five samples for a total of 2560 samples per PPM frame. Each data set contained 76 PPM frames.

In both cases, the sampling time was constant at 40 ns. Hence, 32 PPM was transmitted at $9.7656 \cdot 10^4$ frames per second or $4.8828 \cdot 10^5$ bits per second. In the case of 256 PPM, transmission occurred at $9.7656 \cdot 10^3$ frames per second or 78125 bits per second.

Two types of detectors were used in the experiments:

1. A single detection element in the focal plane. This is called the “single detector.”
2. A focal plane array. This is called the “FPA.”

The single detector featured a higher quality amplifier circuit and notably lower thermal noise than the FPA detector did, resulting in much better BER performance as a function of K_s , which is the number of photons per PPM pulse, or “photons per bit,” which is K_s divided by five bits per symbol for 32 PPM or K_s divided by eight bits per symbol for 256 PPM.

A number of difficulties were experienced with the gathering of experimental data:

1. Each data set requires a large amount of time, typically 10 minutes, to gather, due to equipment limitations of the GageScope software and the computer

hardware used in the data gathering process. Since it takes 10 minutes to gather only 76 symbols (256 PPM), it is difficult to obtain good BER data at higher SNRs due to the scarcity of error events.

2. For a 256 PPM data set of 76 symbols, the lowest BER in a given set is only $1.32 \cdot 10^{-2}$. The only way to obtain lower BER points for a BER plot would be to gather many sets of data at the same (or very nearly the same) SNR and to average the achieved bit error rates. It would take one error out of 13 data sets to obtain an experimental BER of approximately 10^{-3} , and such a data point would be unreliable since only one error event would have occurred. To obtain 20 error events at this BER, one would need to take approximately 263 data sets, which is completely infeasible given the limitations of the equipment in the laboratory.
3. Computation of the SNR required direct analysis of the data. The focal plane array used has a very small fill factor less than 20%, with the result that any changes in the focal plane field distribution would result in an unpredictable total energy striking the detector elements. Since the low fill factor results in a large portion of the signal striking “dead regions” in the detector focal plane, the actual photons received would be significantly less than what one would expect from power meter measurements. Signal-to-noise ratios were still obtained by analysis of the output data from the FPA or from the single detector, depending on which device was used. The methods used to estimate the SNR from the output data were given in Chapter 3.

4. Since only 76 PPM symbols were present in each 256 PPM data set, and since LMS convergence at moderate-to-high SNR is dominated by the signal slots (see Section 5.3), only a limited amount of time was available for LMS convergence. Fortunately, the LMS algorithm converges rapidly, leading to good overall performance even with short data sets.

Items 1 and 2 above impose a limit on the accuracy of the experimental BER curves at high SNR where bit error rates are low. Item 3 necessitated the SNR estimation methods described in Chapter 3.

6.1.1 Operation of the software receiver

6.1.1.1 Overview

The software receiver is designed to accept intermediate frequency (IF) samples. It performs a mixing operation with a local oscillator to convert these samples to baseband, performs a frequency centering operation to bring the resulting signal as close to DC as possible, and performs coherent combining via LMS along with symbol demodulation. A block diagram of this receiver is shown below:

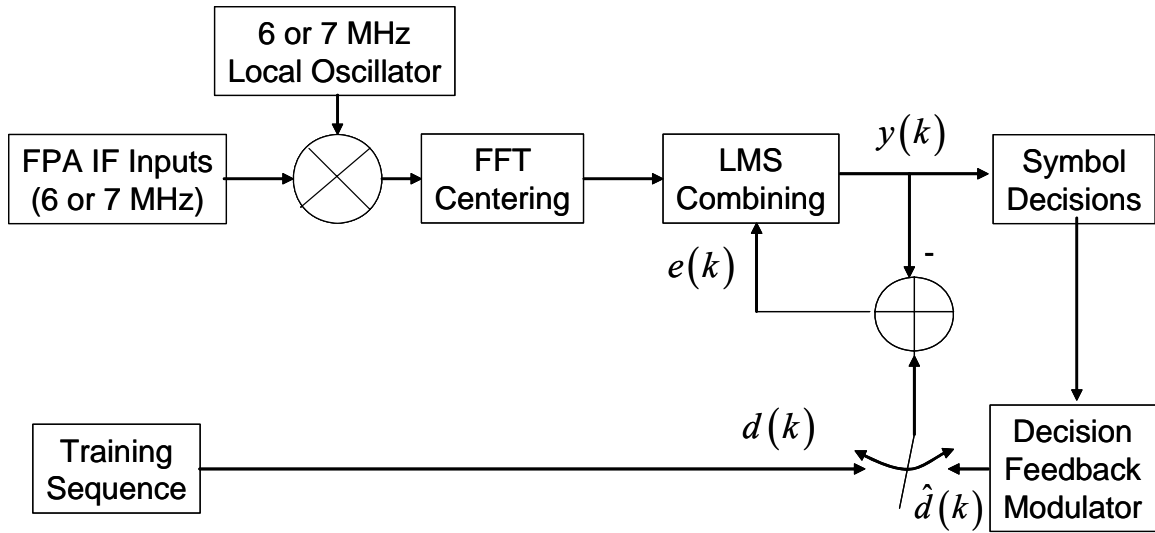


Figure 62. Block diagram of the software receiver.

The operation of this receiver will be described in the next few sections.

6.1.1.2 Correction for frequency offset using FFT centering

Since the exact beatnote frequency that results from the difference in frequencies between the signal laser and the local oscillator is not precisely known, there often exists a small frequency offset from DC after the LO mixing stage. In order to insure downconversion to baseband, the following operations are performed:

1. An FFT of the strongest individual channel signal is taken.
2. The peak of the magnitude of this FFT is located.
3. The original FFT is circularly shifted in order to center the peak at DC.
4. The circularly shifted FFT is then processed with an inverse FFT to yield a new, frequency-corrected signal.

5. All of the FPA channels (in the FPA case) are frequency shifted by an amount that is determined by the circular shift used for the strongest channel above.

The FFT of each of the remaining channels is taken, circularly shifted, and inverse transformed back into the time domain.

The above procedure provides a simple but effective way to avoid rapid phase ramping caused by imperfect downconversion. However, it can be improved in future work. The following improvement has not been implemented at the time of this writing: When the LMS algorithm has converged, use the FFT of the LMS output to determine the amount of the circular shift for each of the input channels.

Continuously update the phase correction using a time-domain windowing technique. The present system simply performs a single correction step over the entire data set. This method does not yield the highly idealized phase tracking of a PLL. If there is a slow phase drift near DC, this method will not, in general, track it out as a PLL would.

6.1.1.3 Decision feedback system

Although most mathematical analyses of LMS in the literature assume the presence of a known desired signal, it is necessary in general to use decision feedback. Since the objective of a communications system is that of demodulating an unknown data stream, it is impossible to rely entirely on a known sequence. Accordingly, decision feedback plays a well-known and important role in adaptive systems, with decision feedback equalizers being one of the most popular and important applications of decision feedback (Proakis, 2000), (Haykin, 1996).

6.1.1.4 Decision feedback operation

The decision feedback system's operation for 256 PPM is described in detail here, and the description for 32 PPM will be given later. The modulation used has the following parameters:

1. PPM order: 256.
2. Slots per PPM frame including dead time: 512.
3. Samples per slot: 5.
4. Samples per PPM frame: 256.

During the first ten symbols, the system operates in acquisition mode. A known reference sequence, obtained by taking advantage of the demodulator reference output, is used to compute the sequence of target samples $d(k)$. Since there are 2560 samples per frame, and since there are ten frames (i.e., PPM symbols) used for acquisition, it is necessary to use the modulator reference over the first 25,600 samples to allow the weights to converge to a point where demodulator output decisions will be reliable. During this period, the switch shown in Figure 62 is set to the actual modulator reference sequence $d(k)$.

Once the first 10 symbols are finished, the switch shown in Figure 62 is set to $\hat{d}(k)$, which is the output of the decision feedback modulator. Here, symbol decisions from the demodulator are used to drive an internal PPM modulator in the receiver. The output modulator sequence is $\hat{d}(k)$, an estimate of the true sequence $d(k)$. This sequence

is used as the LMS training sequence. For each PPM frame, the decision feedback system performs the following steps:

1. Demodulate the output of the LMS combiner in the PPM demodulator. This results in a symbol decision.
2. Send the resulting symbol decision to the decision feedback modulator. This generates a noiseless PPM frame (noiseless in the sense of being free from Gaussian noise, although this frame would be in error if the wrong decision was made at the PPM demodulator).
3. Using the reference frame generated by the internal modulator, which is 2560 samples long, run LMS training over the entire duration of the current PPM frame in order to update the weights.

The above procedure is used for each PPM symbol after the first ten symbols.

The process for 32 PPM is similar. The only differences are:

1. A total of 25 symbols are used for acquisition. This is possible because a data set contains a total of 779 symbols.
2. Since there are 32 slots with eight samples each, there are 256 samples per frame.

6.1.2 Synchronization issues in the optically coherent PPM system

PPM is a modulation that requires accurate knowledge of both slot and frame boundaries. Since these are temporal boundaries, any errors in timing will result in demodulation errors. Such errors are unacceptable since the experimental setup is used to obtain bit error rate (BER) data to compare experimental outcomes with theoretical predictions.

The subject of PPM symbol timing recovery is beyond the scope of this thesis, and this problem has been covered in the open literature in Sun and Davidson, (1990) and Vilmrotter, Rodemich and Tan, (1986) among other sources. However, recovery of symbol and slot timing information in the receiver remains an important problem: reliable PPM demodulation is necessary for the computation of bit error rates in the laboratory.

The experimental setup in the laboratory did suffer from a timing offset between the transmitter and the receiver along with a slowly varying timing drift. Since the reference output of the PPM modulator was readily available, timing control in the receiver was a matter of examining the modulator reference output and using it to perform both slot and symbol synchronization. In this thesis, the terms “PPM symbol,” “PPM frame,” “symbol,” and “frame” are all used interchangeably. This timing error is illustrated in Figure 63.

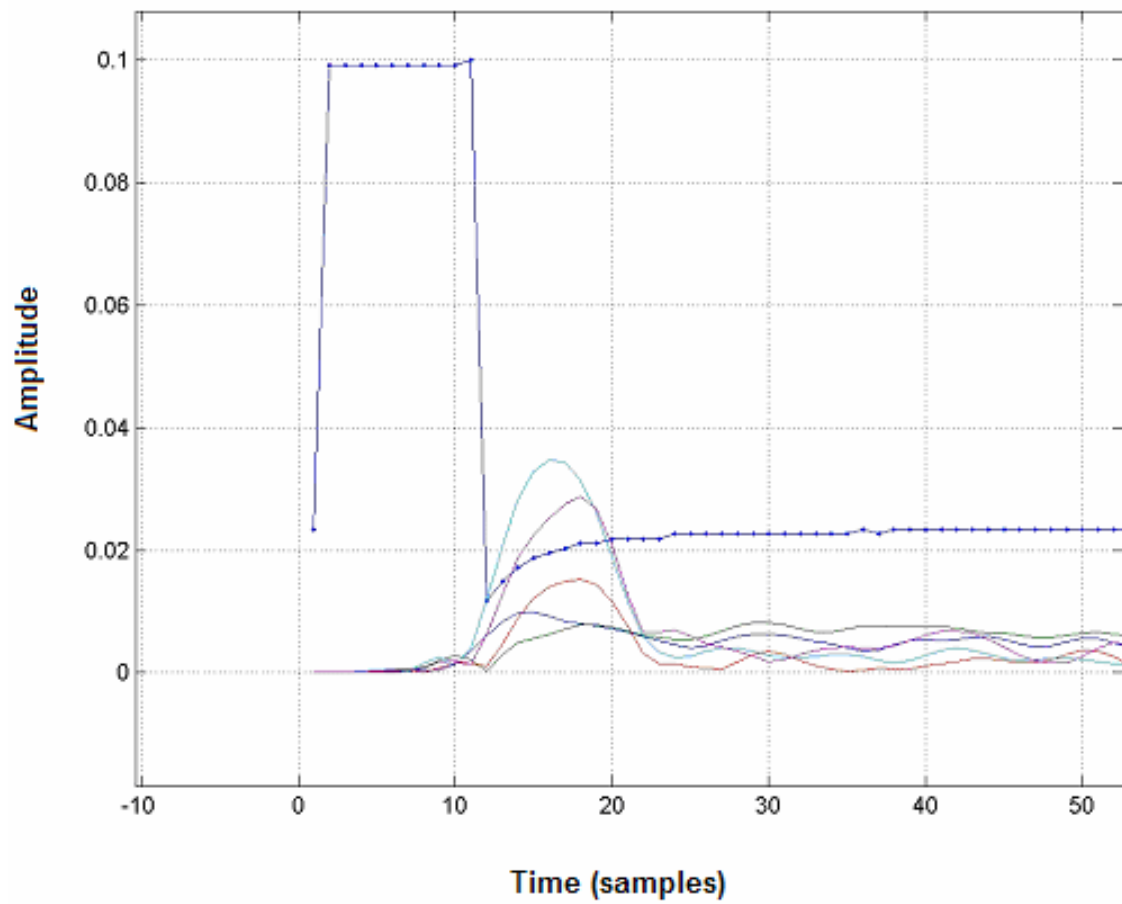


Figure 63. Example of time misalignment prior to correction. There are five FPA channel signals that suffer a delay of approximately 10 samples relative to the modulator reference pulse.

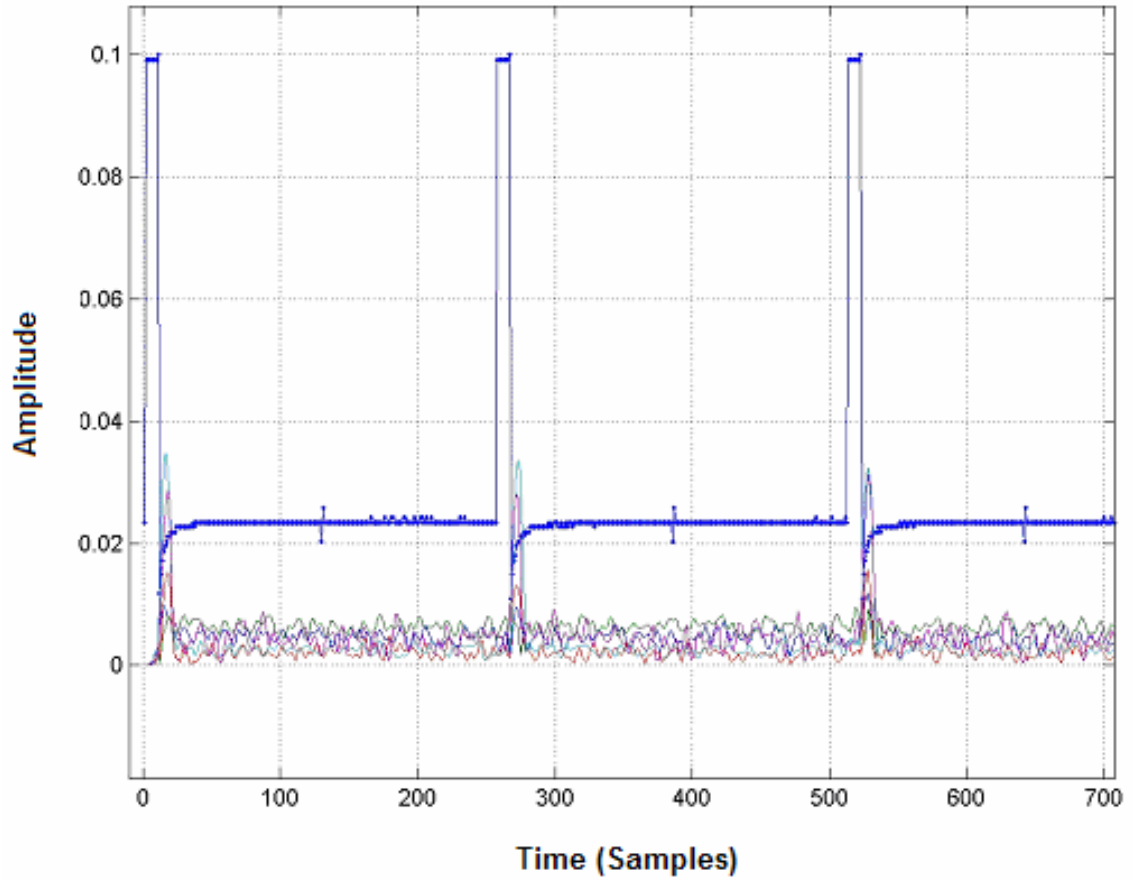


Figure 64. Time misalignment between modulator reference pulses and receiver prior to correction. The misalignment only changes slowly over time.

The approach taken in the software receiver does not attempt to implement the slot and symbol synchronization approaches present in the literature. Instead, the rising edge of the modulator reference signal is used for synchronization purposes. Since the modulator reference output has an amplitude of 0.10 V, the software receiver can reliably detect the rising edge of the modulator pulse, and rising edge detection is the key to the software receiver's synchronization subsystem.

First, after taking a large number of data sets, it was found that the typical delay between the modulator pulse and the signal pulse observed at the receiver is 10 or 11 samples, with slight drifts of one or two samples over time. The steps taken to recover timing information at the receiver are described here.

First, the mean symbol timing error, in samples, was computed from the data. The peaks of the received signal pulses without any attenuation are located, and their locations are compared to the locations of the centers of the rectangular modulator pulses. The average distance of 10 samples was calculated from this.

Although the average distance from the modulator reference pulse centers to the receiver pulse centers is 10 samples, this distance changes during the course of data acquisition. The clocks in the modulator and in the receiver are not synchronized in the laboratory, and they would not be synchronized in a realistic deep space communications link either. To compensate for this timing drift, it is necessary to monitor the locations of the rising edges of the modulator reference pulses. Whenever a modulator reference pulse arrives, the following steps are taken:

1. The rising edge of the pulse is always assumed to denote the start of a new slot.
2. The location of the rising edge, which is really the index of the first sample of the modulator reference stream exceeding a fixed threshold, is compared to the slot boundaries for the current frame.
3. The difference between the location of the rising edge and the closest slot boundary is determined.

4. The PPM data frame window is shifted so that the slot boundary closest to the rising edge coincides exactly with the rising edge.

This strategy is meant only for the laboratory and is not a viable solution in an actual deep space communications system. Most of the misalignment of the pulses is caused by a fixed 10 sample delay between the modulator reference pulses and the receiver pulses, and the rising edge based method for maintaining slot alignment plays a less important role, usually correct one or at the most two samples of drift. However, in the case of 256 PPM with only five samples per slot, a two sample drift in timing corresponds to 40% of the slot width. Hence, the rising edge detection capability to correct for timing drift is necessary. Timing correction cases are illustrated in Figures 59 and 60.

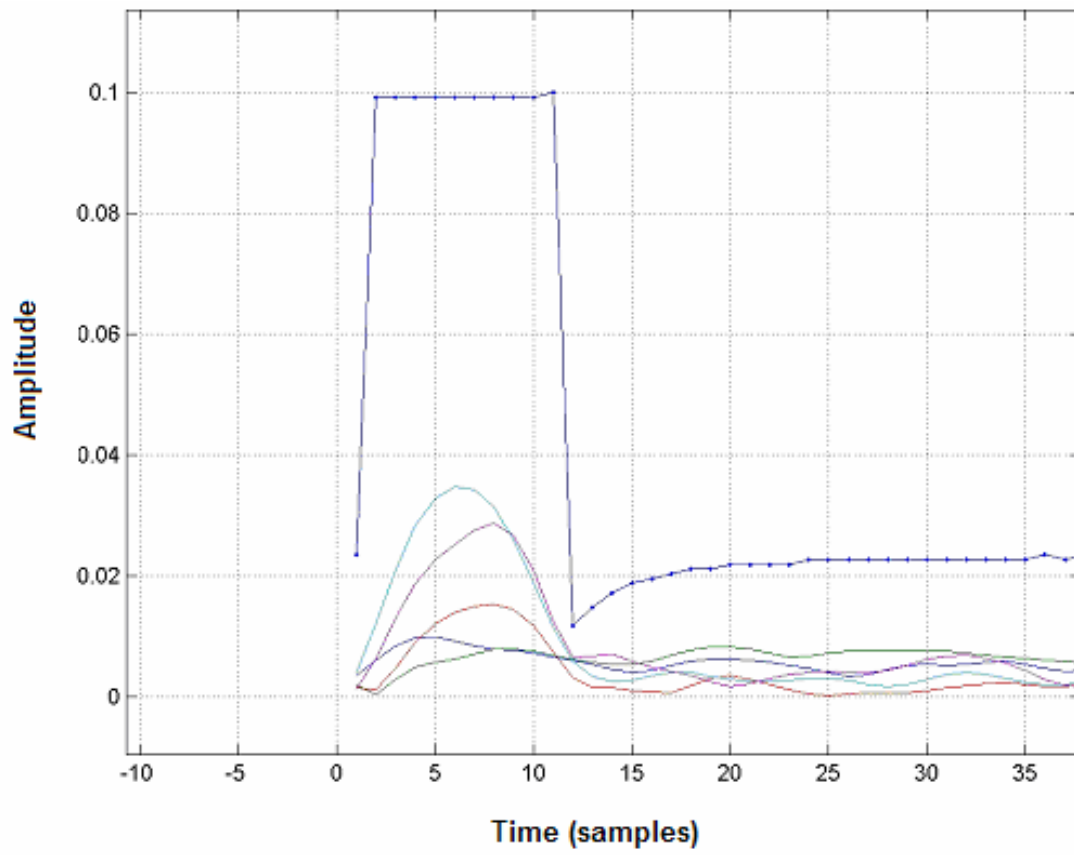


Figure 65. Time-corrected pulse alignment, corresponding to the case shown in Figure 63.

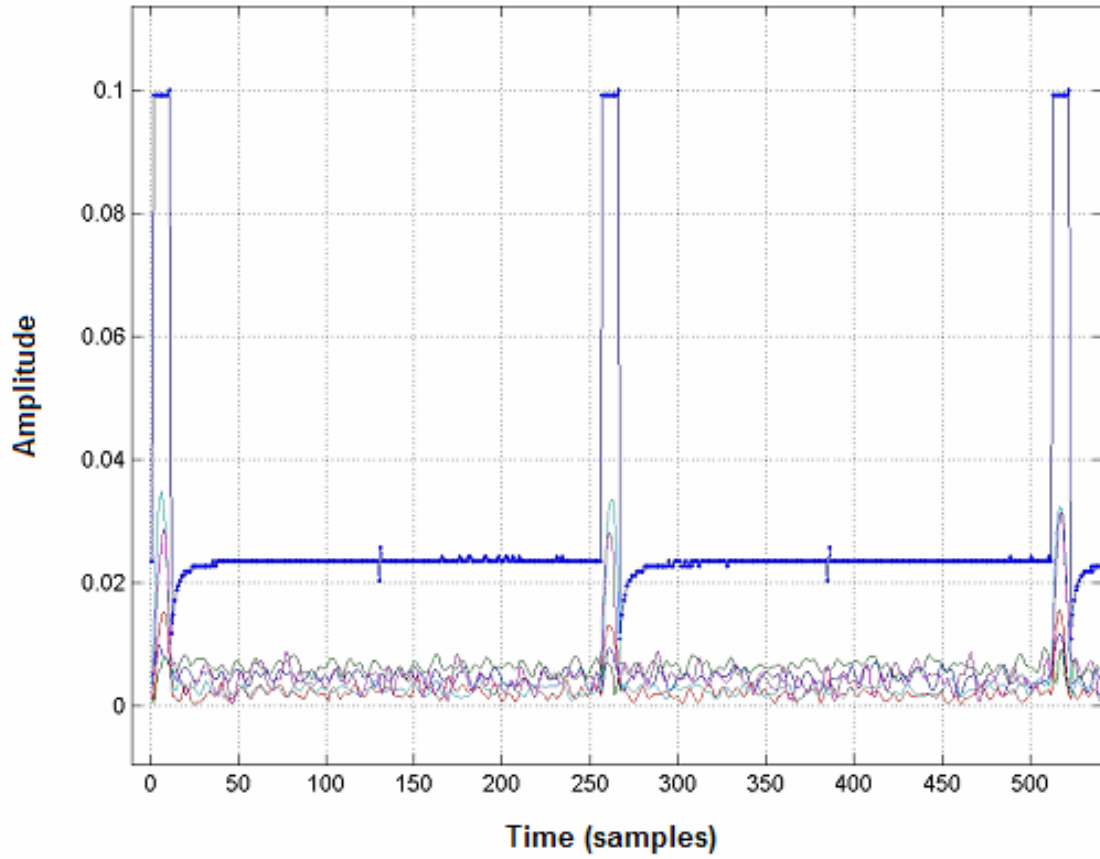


Figure 66. Time corrected pulse alignment, corresponding to the case shown in Figure 64.

6.2 Evaluation of experimental system performance and comparison with theory

6.2.1 BER performance bounds

The lower bound on achievable BER is given by the theoretical BER curve for the no leakage case derived in Chapter 3. The upper bound was also computed, using measured

leakage statistics, in the same chapter. In each of the single detector BER plots, four items are shown

1. The experimental data points.
2. The BER trend line fitted to these data points.
3. The BER upper bound curve based on measured leakage data.
4. The BER lower bound curve based on the no leakage case.

The BER trend line was fitted to experimental data. As such, it will tend to be less reliable outside of the regions where many points were gathered. For this reason, the BER trend line may, in certain instances, fail to agree with the upper and lower bounds at extremely low SNR or at extremely low values of K_s where two effects occur. First, since SNR is very low the reliability of the SNR estimate is also low, making reliable curve fitting difficult in the low SNR region. Second, only a few points were gathered at very low combined SNR, and the scarcity of points in this region will also affect the reliability of trend lines in this area. Near 0 dB SNR, a trend line may therefore fail to exhibit expected behavior and may even cross above or below the upper or lower bound lines, respectively. These effects are limited, but they still appear in some of the experimental data plots.

In each of the FPA plots, the same quantities are plotted for the individual channels. Additionally, the combined channel actual performance is plotted along with the theoretical performance of a single detector system without leakage. In all of the FPA plots, the x-axis represents combined SNR, combined K_s , or combined photons per bit. Hence, the single channel BER data are themselves plotted against the combined

channel's maximum achievable SNR, although the lower bound curves for the single channels are computed using the theoretical expressions for the case without leakage.

The goal of the FPA is to coherently combine the signals from the individual channels in order to achieve the same, or at least nearly the same, performance as an ideal single detector system without any atmospheric turbulence (i.e., with all of the photons reaching the single detector). In the ideal, shot noise limited case, the SNR is given by

$$\text{SNR} = 10 \log_{10} \left(\frac{E_s}{N_0} \right) = 10 \log_{10} (2K_s) \quad (6.2.1)$$

where K_s is the number of photons per PPM pulse. In a multiple channel system, let $K_{s,i}$ denote the number of photons per pulse in the i^{th} FPA channel. The total number of photons per pulse is

$$K_s = \sum_{i=1}^N K_{s,i} \quad (6.2.2)$$

for an N -channel FPA system. Whether or not the system is shot noise limited, the total array SNR, expressed as a ratio and not in decibels, is

$$\text{SNR}_{\text{total}} = \sum_{i=1}^N \text{SNR}_i \quad (6.2.3)$$

where SNR_i is the SNR of the i^{th} channel. This formula agrees entirely with (6.2.2) as well.

If the FPA receiver achieves the same BER performance as a function of total SNR as would have been achieved by a single detector with the same total SNR, then the FPA receiver is successfully combining the signal energies to achieve the same

performance as a single detector receiving the same SNR all in one channel. In this case, the FPA receiver has effectively combined the signal over its multiple elements to regenerate the signal that would have been received directly with a single detector with the same noise characteristics if no phase distortion had been present to spread the energy over the area of the FPA. In that case, the FPA receiver has effectively compensated for the spreading of the signal in the detector plane by recombining the scattered signal energy.

To determine the extent to which the FPA successfully recombines scattered signal energy, the BER of the combined channel is plotted as a function of the total SNR. A theoretical curve that illustrates BER performance as a function of total SNR for an ideal single-channel detector that captures all of the received signal energy without any spreading of the signal due to atmospheric phase is also drawn. The closer the FPA's performance lies to the theoretical curve, the closer the FPA is to combating the effects of atmospheric turbulence and achieving the same performance as that of an optical channel in which no atmospheric phase distortion is present.

6.2.2 Results with the single detector

Experiments involving a single detector were performed using both 32 PPM and 256 PPM modulation. The purpose of these experiments was to illustrate nearly shot-noise-limited performance (shot noise 5 dB above thermal noise). It was verified at the laboratory that the photocurrent noise was due to shot noise. The procedure that was followed was to measure the relationship between electronic noise power and optical power. The relationship was linear. This stems from the fact that the only fluctuations in

the photocurrent due to a coherent state will be shot noise fluctuations. Although the shot noise due to the local oscillator was greater than the thermal noise component, thermal noise was still present. Hence, the results presented here are not truly shot-noise limited.

The relationship between the SNR ratio $\frac{E_s}{N_0}$, expressed as a ratio and not in decibels, and

the number of photons per pulse, was

$$\frac{E_s}{N_0} = 1.52K_s \quad (6.2.4)$$

The first set of BER curves covers the case of random 256 PPM. Figure 67 illustrates BER performance as a function of K_s , which is the number of photons per pulse. The experimental points lie close to the theoretical BER performance line, and the trend line fitted to the experimental points is seen to lie very close to the theoretical bound. The upper bound in this case, computed based on leakage statistics, is seen to be pessimistic. In Figure 68, BER is plotted against photons per bit. Each 256 PPM symbol carries 8 bits of information, and the number of photons per bit is $\frac{K_s}{8}$. This figure shows

a bit error rate only slightly above $3 \cdot 10^{-2}$ at 1 photon per bit, illustrating the system's performance in the presence of weak signals. A plot of BER vs. $\frac{E_s}{N_0}$ is shown in Figure

69.

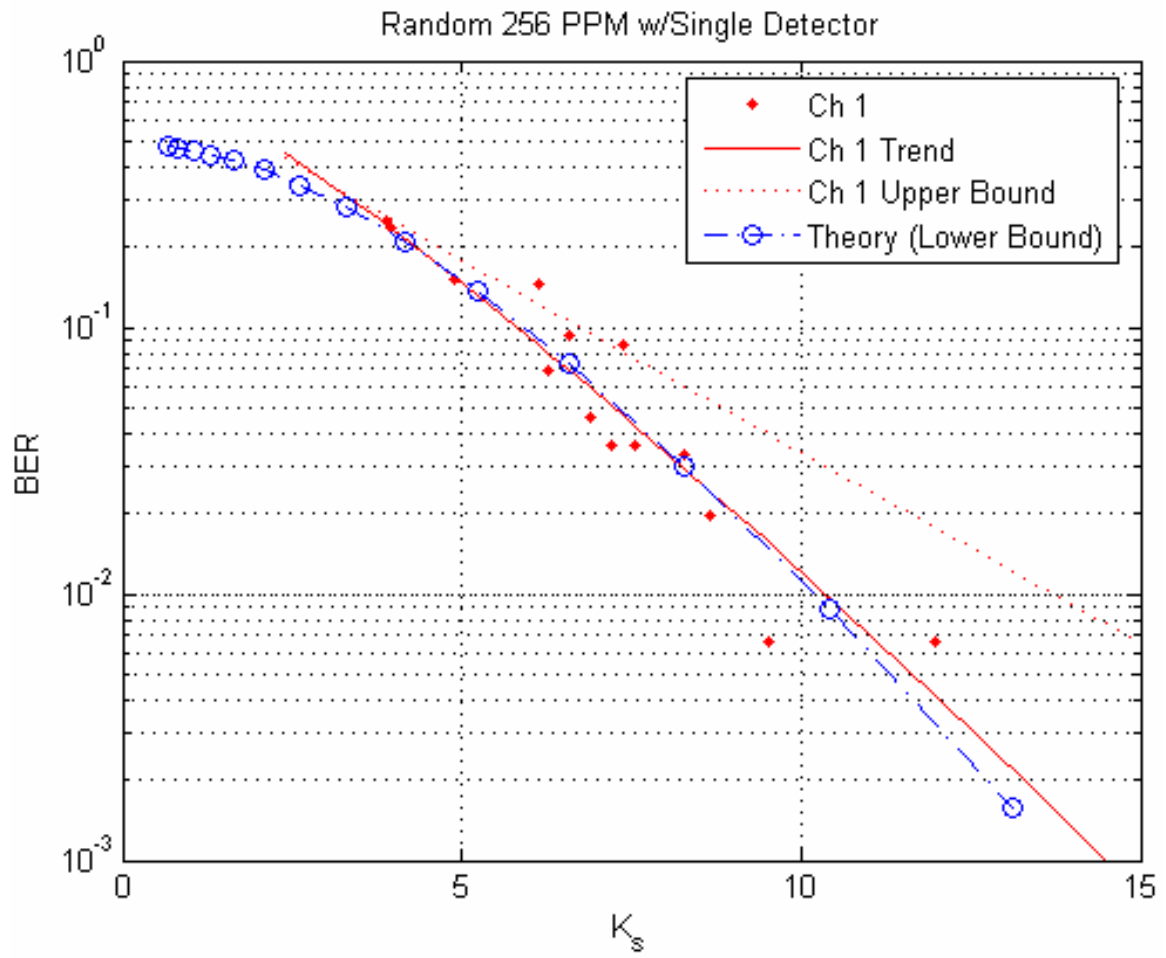


Figure 67. BER vs. photons per pulse with the single detector and 256 PPM.

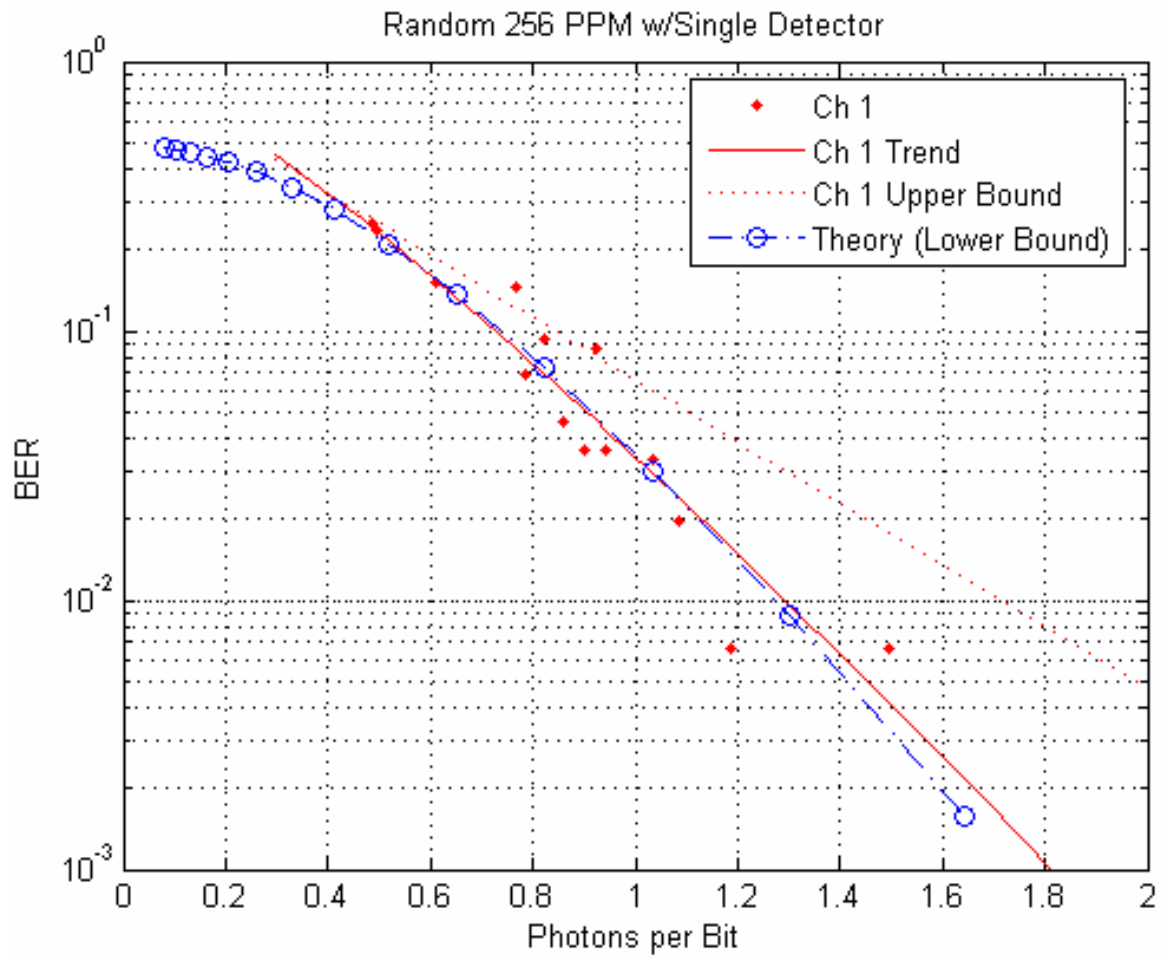


Figure 68. BER vs. photons per bit with the single detector and 256 PPM.

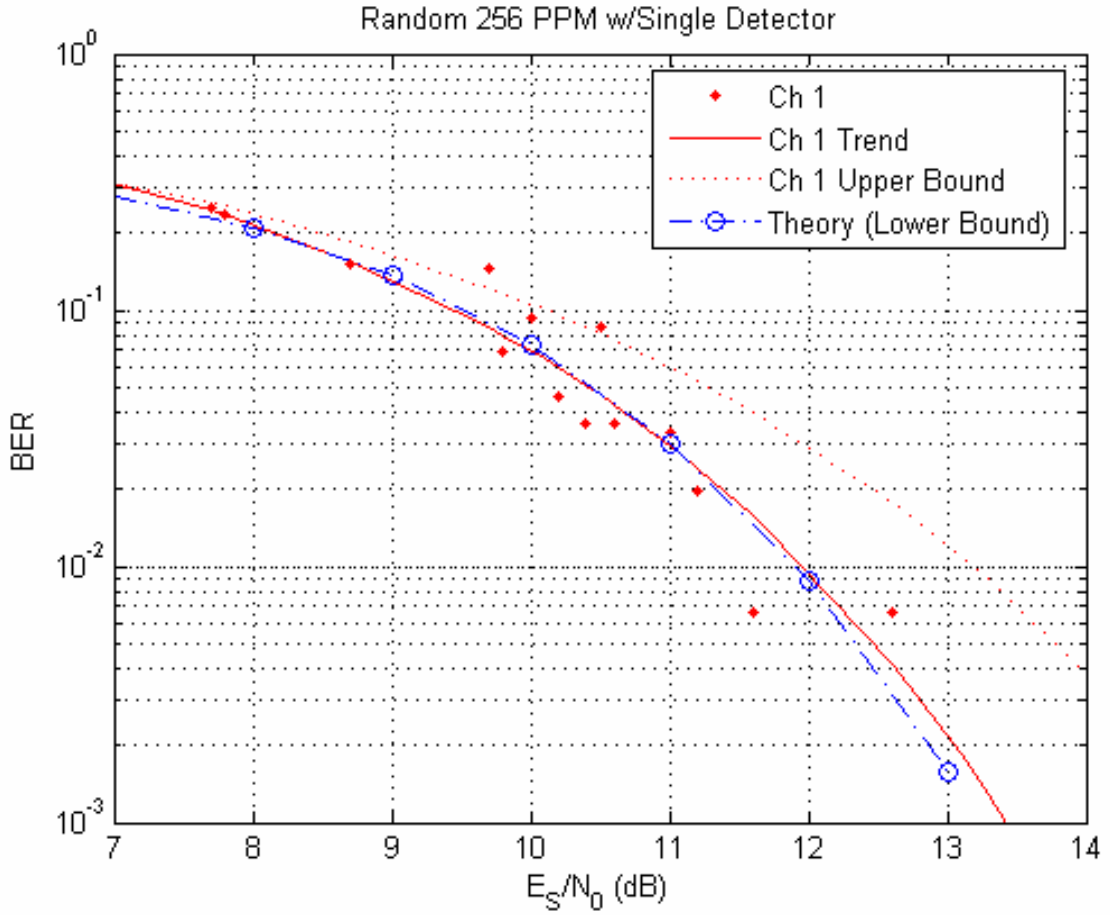


Figure 69. BER vs. E_s/N_0 in dB for the single detector and 256 PPM.

Since all of the noise in the system is assumed to be Gaussian (Gagliardi, 1995), one can convert directly between the signal-to-noise ratio, which is $\frac{E_s}{N_0}$, and the number of photons per pulse K_s provided that one knows the components of the noise spectral density N_0 due to local oscillator shot noise, receiver thermal noise, and background noise. If local oscillator shot noise strongly exceeds the other noise components, then the

noise spectral density will be shot noise dominated, and the system is said to be shot noise limited. In the shot noise limited case

$$\frac{E_s}{N_0} = 2K_s \quad (6.2.5)$$

for optical heterodyning (Gagliardi, 1995). Given that $\frac{E_s}{N_0}$ is available, one could plot hypothetical BER curves that would have been obtained with a shot noise limited receiver. One would have to solve Eq. (6.2.5) for K_s . Using this procedure, the following hypothetical BER curves for a shot-noise limited receiver were obtained.

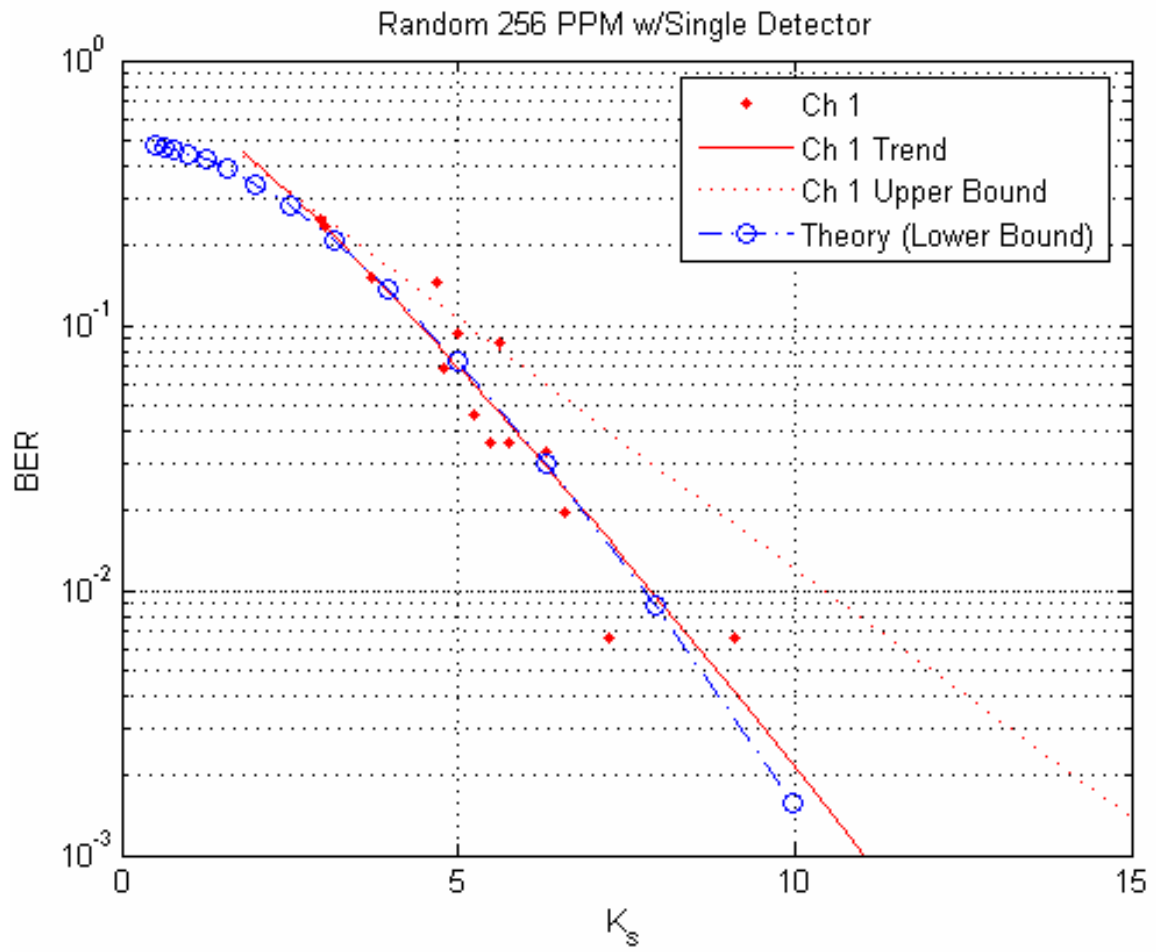


Figure 70. BER vs. K_s assuming a hypothetical shot-noise limited single detector with 256 PPM.

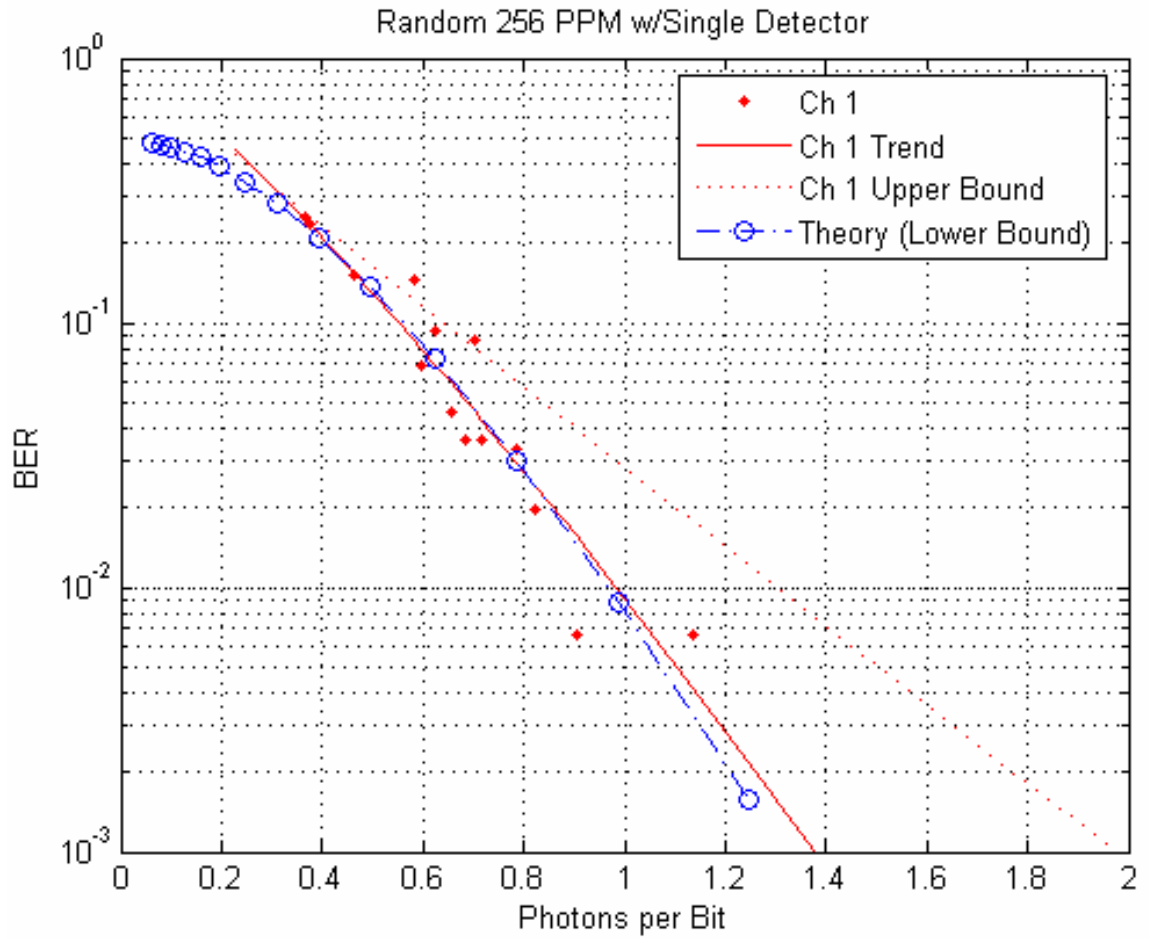


Figure 71. BER vs. photons per bit assuming a hypothetical shot-noise limited single detector with 256 PPM.

Note that BER performance at 1 photon per bit is slightly better than 10^{-2} . A similar set of curves were also generated with fixed 32 PPM.

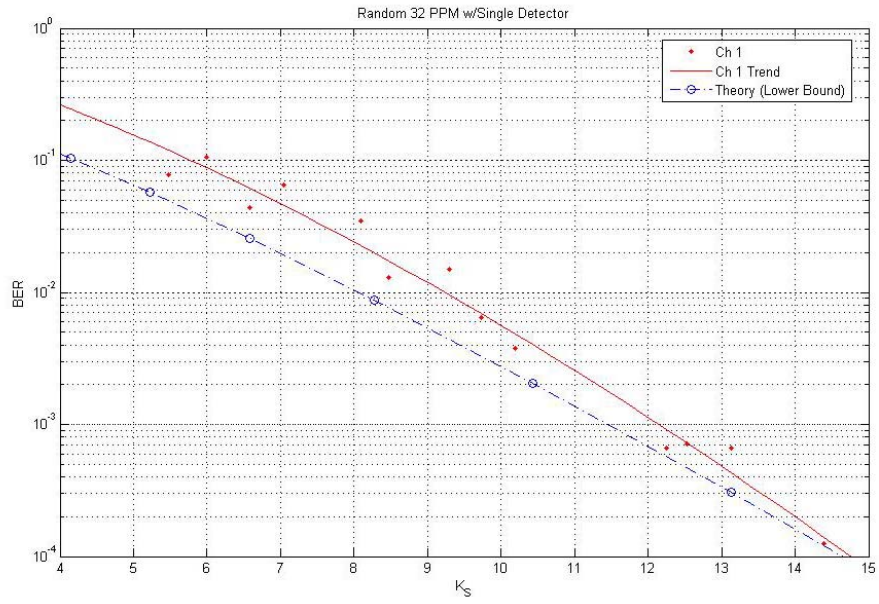


Figure 72. BER vs. K_s with the single detector and 32 PPM.

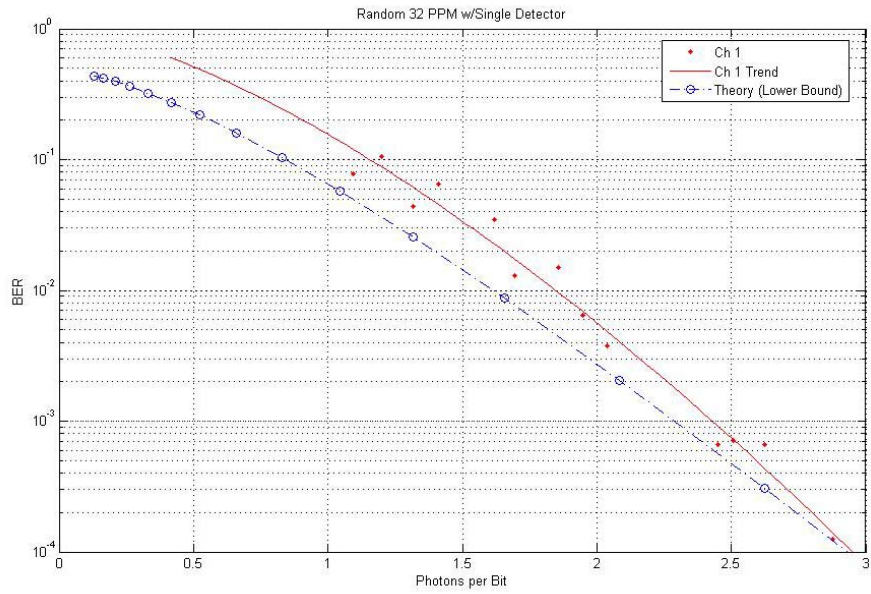


Figure 73. BER vs. photons per bit with the single detector and 32 PPM.

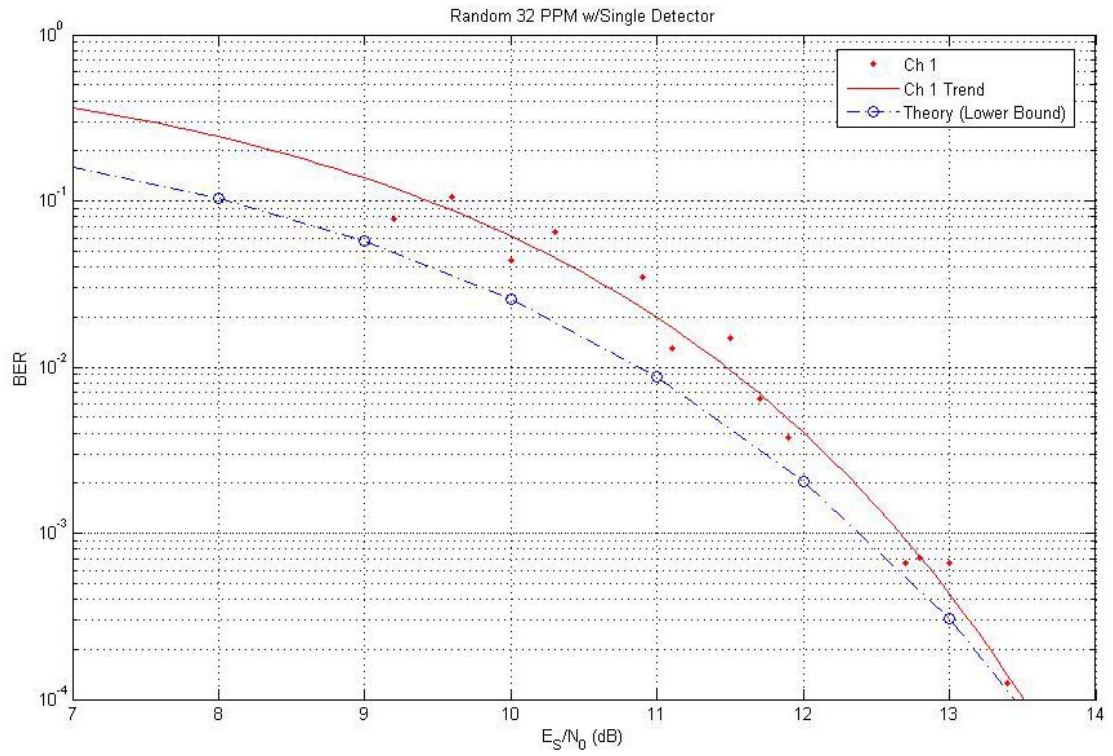


Figure 74. BER vs. E_s/N_0 in dB for 32 PPM with the single detector.

As with the 256 PPM case, the actual relationship between K_s and SNR is given by Eq. (6.2.4). Again, the performance that could have been achieved with a shot noise limited detection system is plotted by solving for K_s in Eq. (6.2.5). The results are shown below.

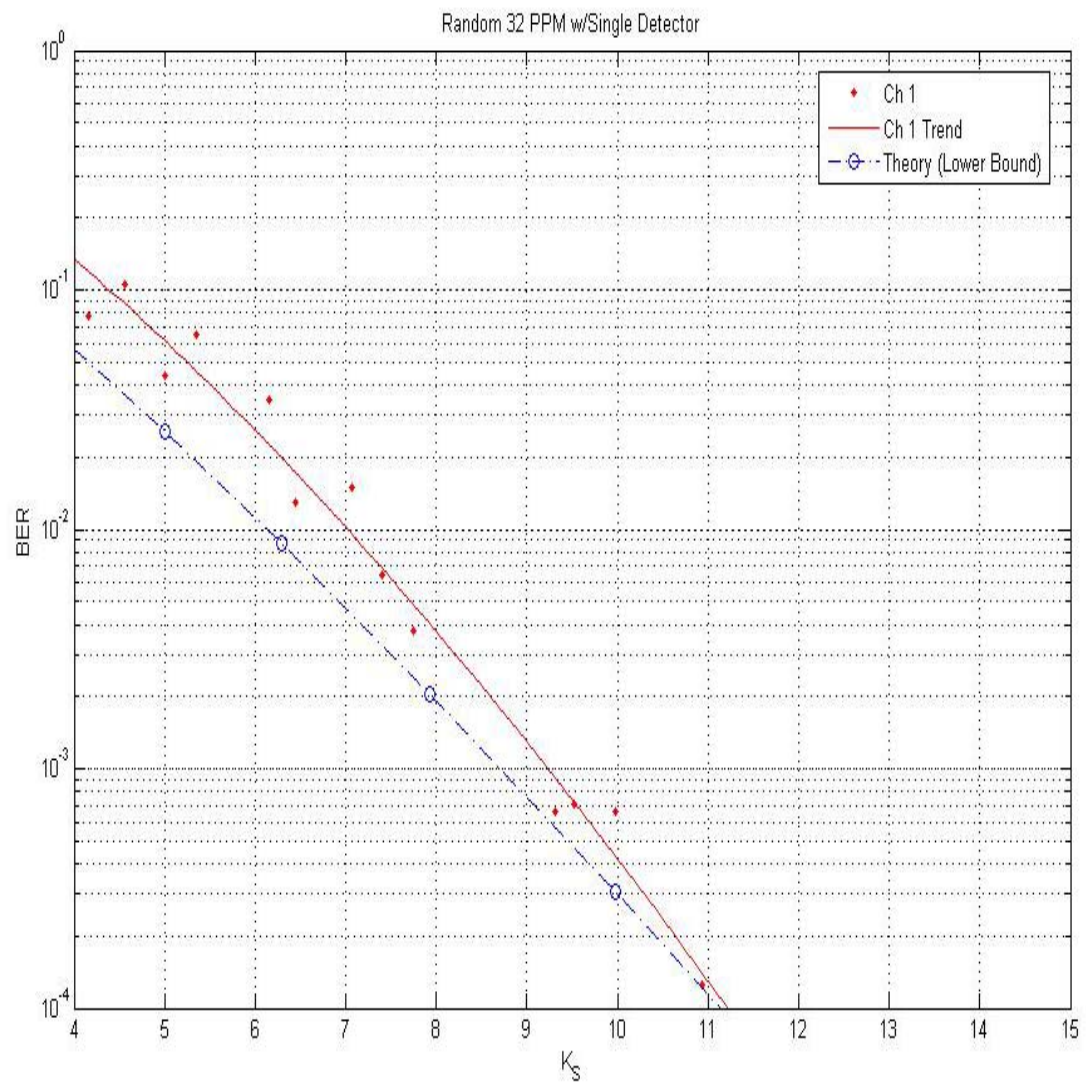


Figure 75. BER vs. K_s assuming a shot-noise-limited single detector and 32 PPM.

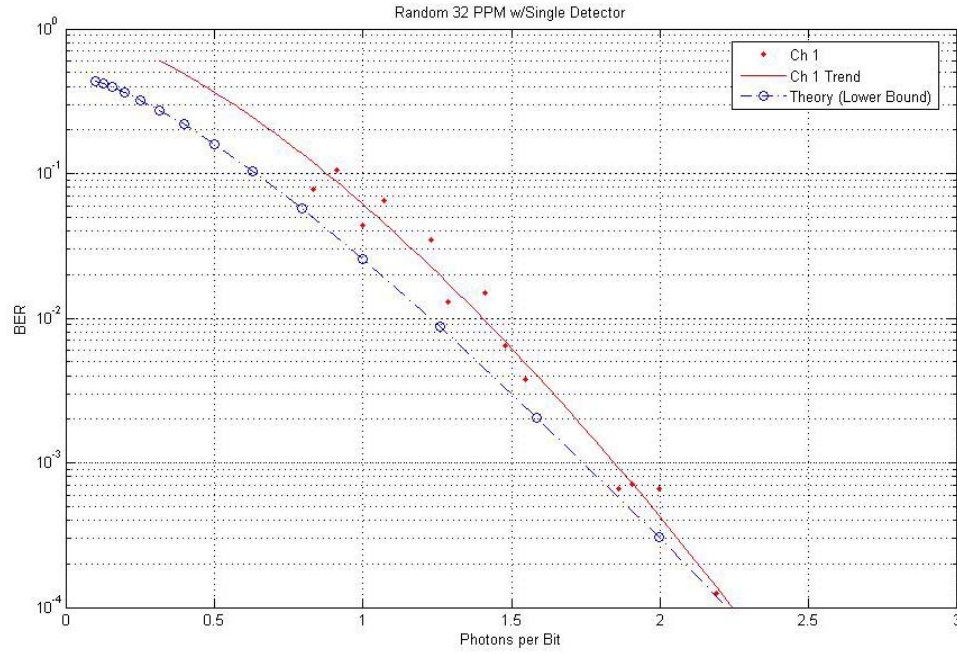


Figure 76. BER vs. photons per bit assuming a shot-noise-limited detector with 32 PPM.

The experimental results using a single detector illustrate the ability to receive telemetry even with signals as weak as 1 photon per bit or weaker. Furthermore, actual single detector performance, while not shot noise limited, still illustrates very good weak signal reception capability.

6.2.3 Results with the focal plane array

The objective here is to demonstrate the ability to coherently combine multiple channels in a focal plane array in order to recover the desired signal. The goal is to come as close as possible to the theoretical BER curve for BER vs. K_s or BER vs. photons per bit. The FPA, by allowing signals from different parts of the detector plane to be summed

coherently, is able to come close to the level of performance that would be achieved with a single detector without any phase distortion in the aperture plane.

The FPA amplifiers introduced a great deal of thermal noise into the system. Performance was not at all shot noise limited, and the Eq. relating SNR to K_s is

$$\frac{E_s}{N_0} = 0.14K_s \quad (6.2.6)$$

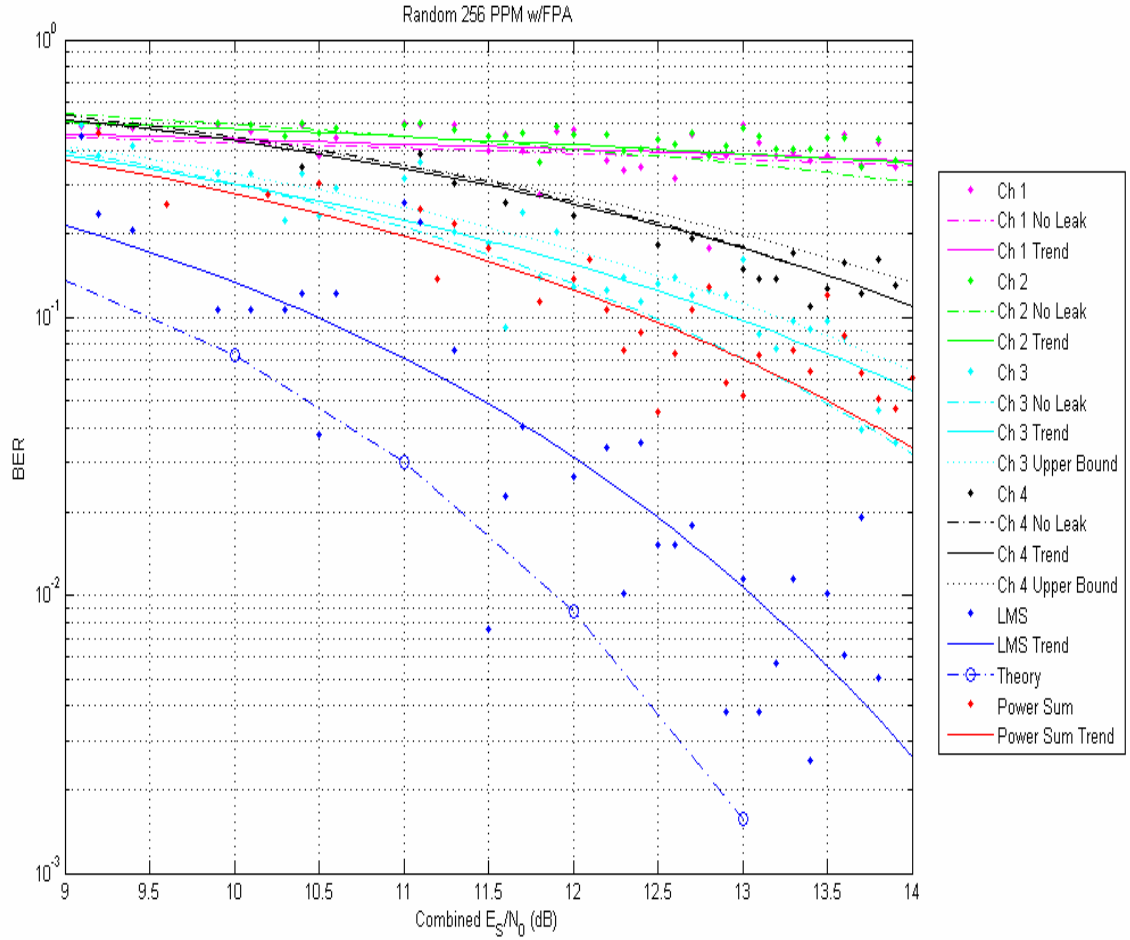


Figure 77. BER vs. E_s/N_0 with the FPA for 256 PPM.

From Figure 77, it is seen that the combined channel is within 1.0 to 1.5 dB of the theoretical curve, indicating that the FPA achieves good combining performance. It should also be noted that the FPA's combined BER curve is almost 3 dB better than the best single FPA channel. This implies that a single detection element under the best of conditions would suffer from a 2.5 to 3.0 dB loss in performance relative to the FPA. In Figure 77, a comparison between the combined channel using the LMS algorithm and the combined channel using only a sum of individual channel powers was made. The sum of powers channel is a reasonable and intuitive performance baseline. It is observed that the BER trend line for the sum of powers channel nearly coincides with the lower bound on BER for channel 3. The sum of powers channel is observed to yield significantly worse performance than the coherently combined channel. For this reason, the sum of powers combined channel will no longer be considered.

The FPA's gain is reasonable. Four channels are being coherently combined. Channels 1 and 2 are very weak and contribute only a small amount to the total SNR. The two stronger individual channels are about 1 dB apart, and combining them coherently would lead to a gain of slightly more than 2.5 dB in theory. Since SNR estimates obtained with the methods of Chapter 3 are themselves random variables, a certain amount of estimation error is to be expected.

Given the SNR expressed as the ratio $\frac{E_s}{N_0}$, one could determine both K_s and the number of photons per bit if the FPA had been capable of shot noise limited operation.

The hypothetical plots of BER vs. K_s and BER vs. photons per bit are shown in Figures 78 and 79.

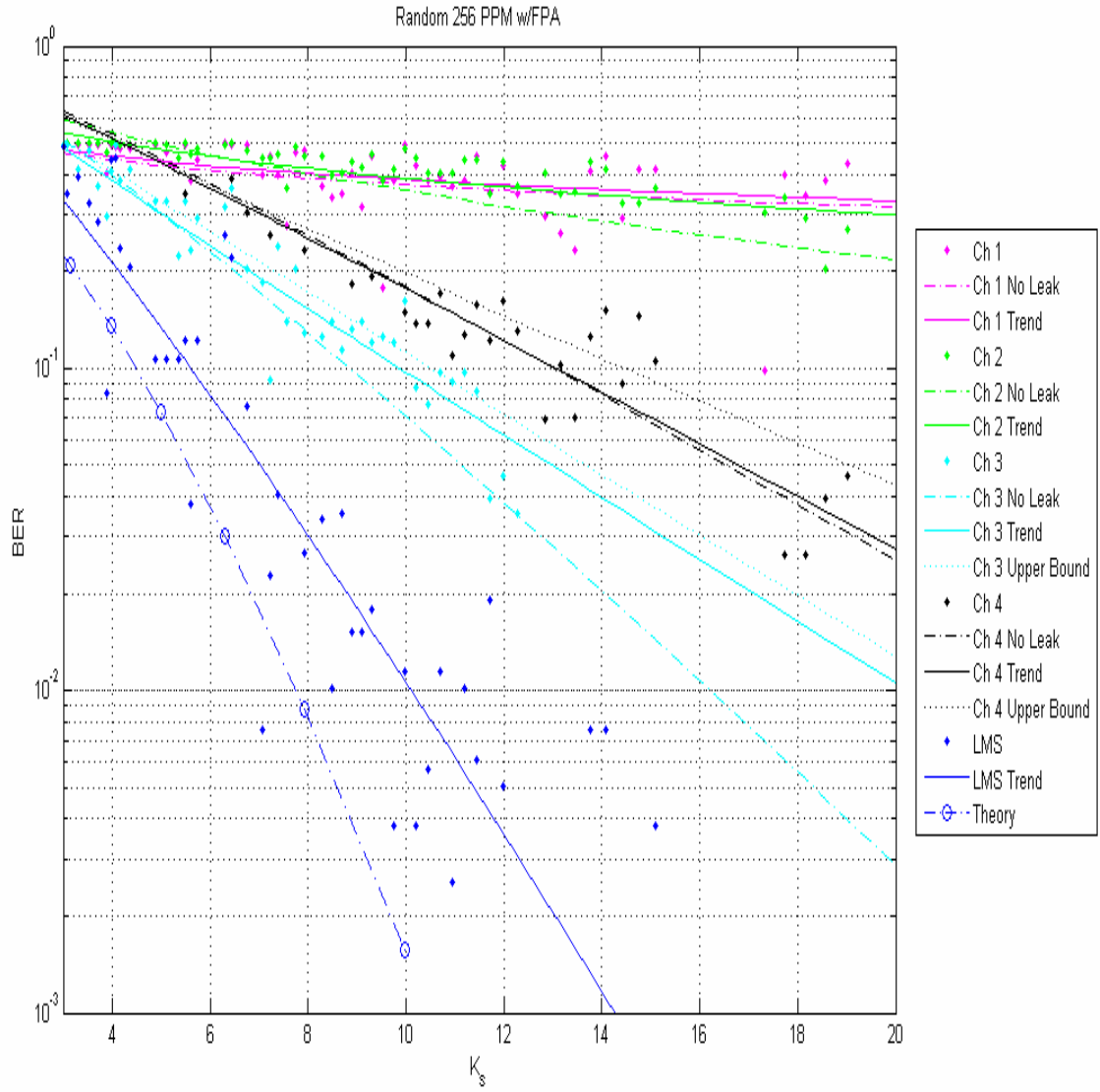


Figure 78. BER vs K_s for a shot noise-limited FPA system with 256 PPM.

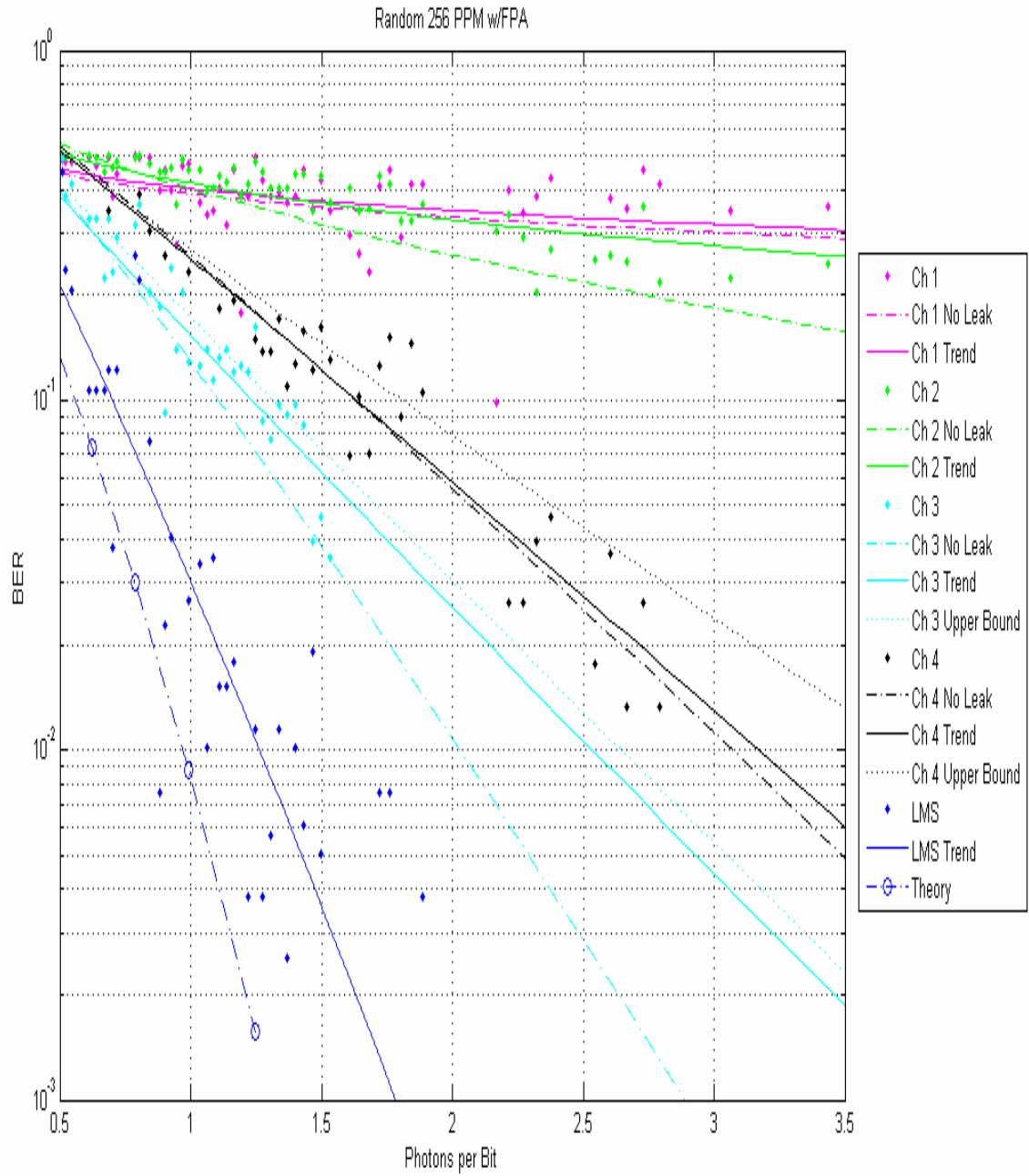


Figure 79. BER vs. photons per bit for a shot-noise-limited FPA system with 256 PPM.

Similarly, one could compute hypothetical FPA performance if the detector elements had the same characteristics as the single detector, and this is done in Figures 80 and 81.

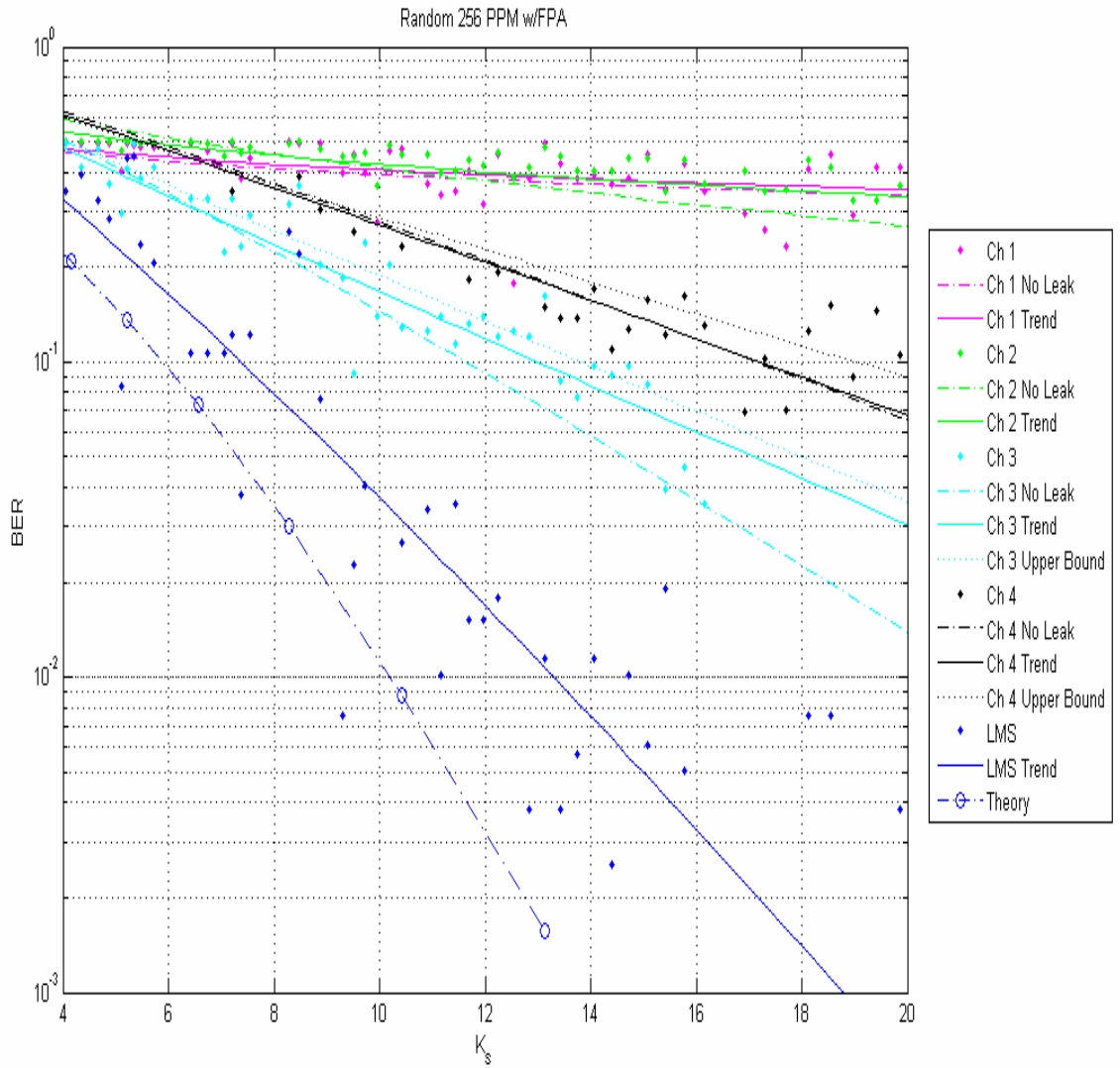


Figure 80. BER vs. K_s assuming an FPA with the same characteristics as the single detector (256 PPM).

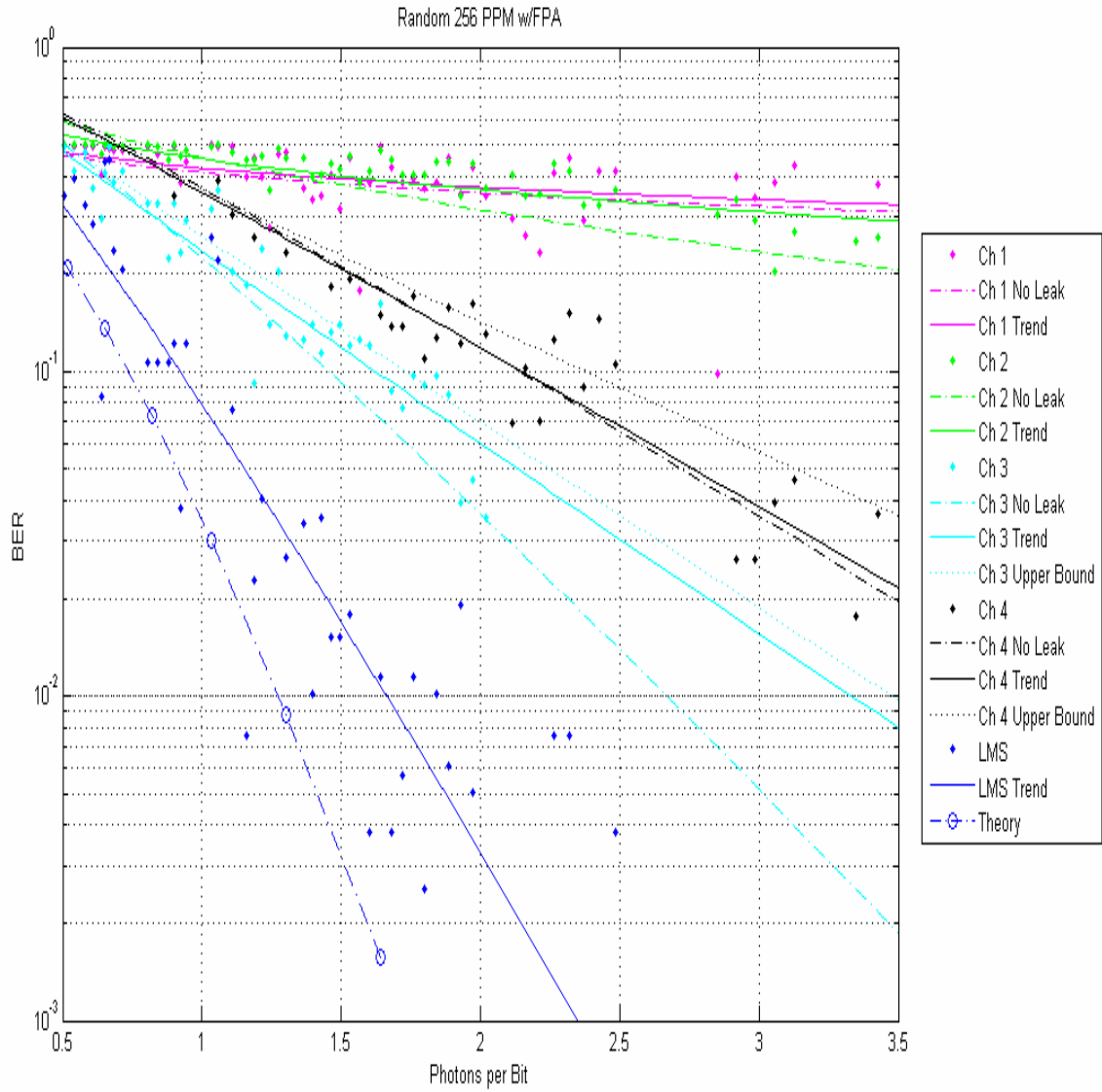


Figure 81. BER vs. photons per bit assuming an FPA with the same characteristics as the single detector (256 PPM).

The preceding figures clearly illustrate the FPA's potential with improved detectors and/or a stronger local oscillator to achieve shot-noise-limited or near-shot-noise-limited BER performance. However, since the actual system suffered from high

levels of thermal noise, it was necessary to determine the actual achieved performance, which is shown in Figures 82 and 83.

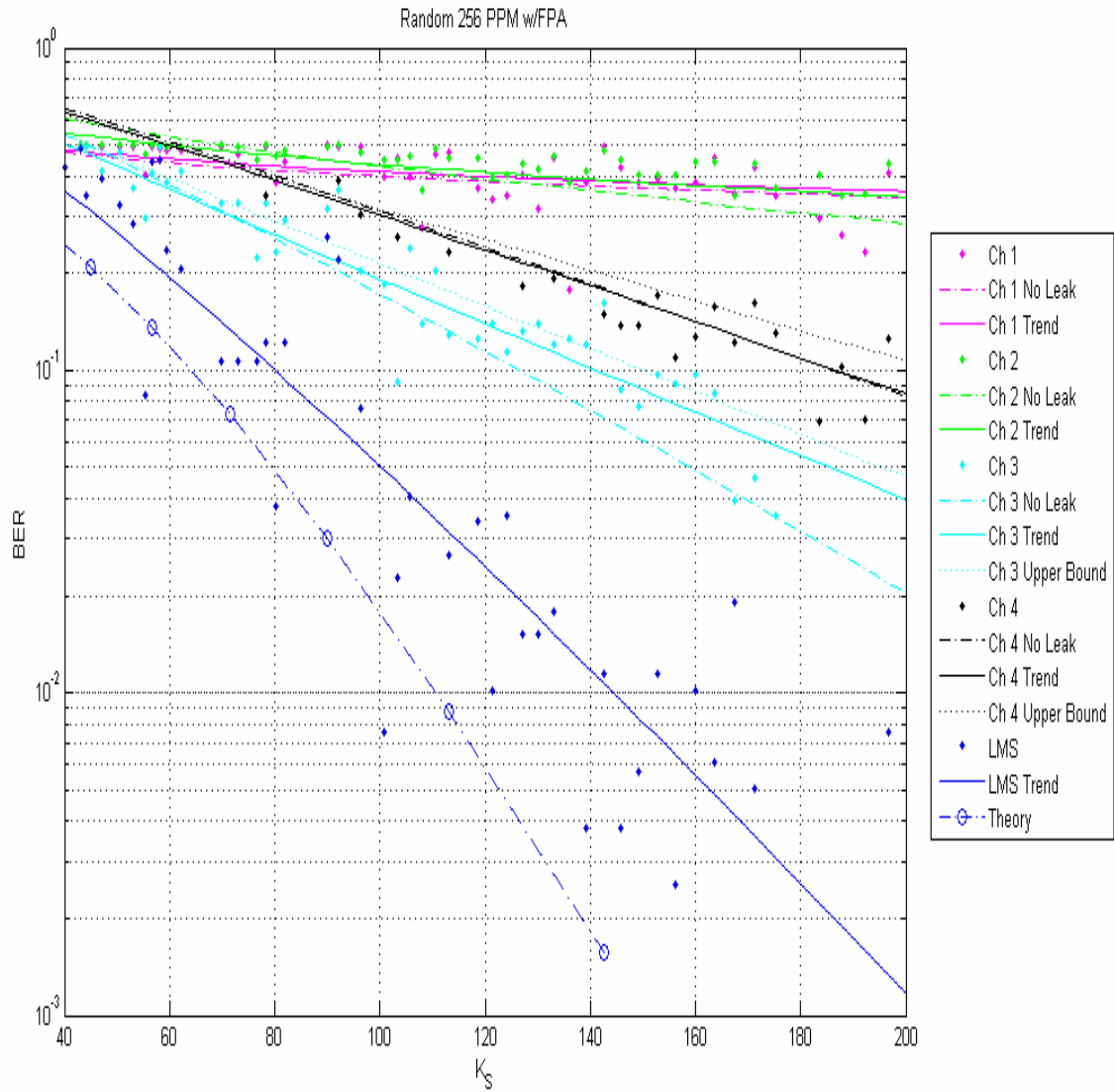


Figure 82. BER vs. K_s for 256 PPM with the FPA.

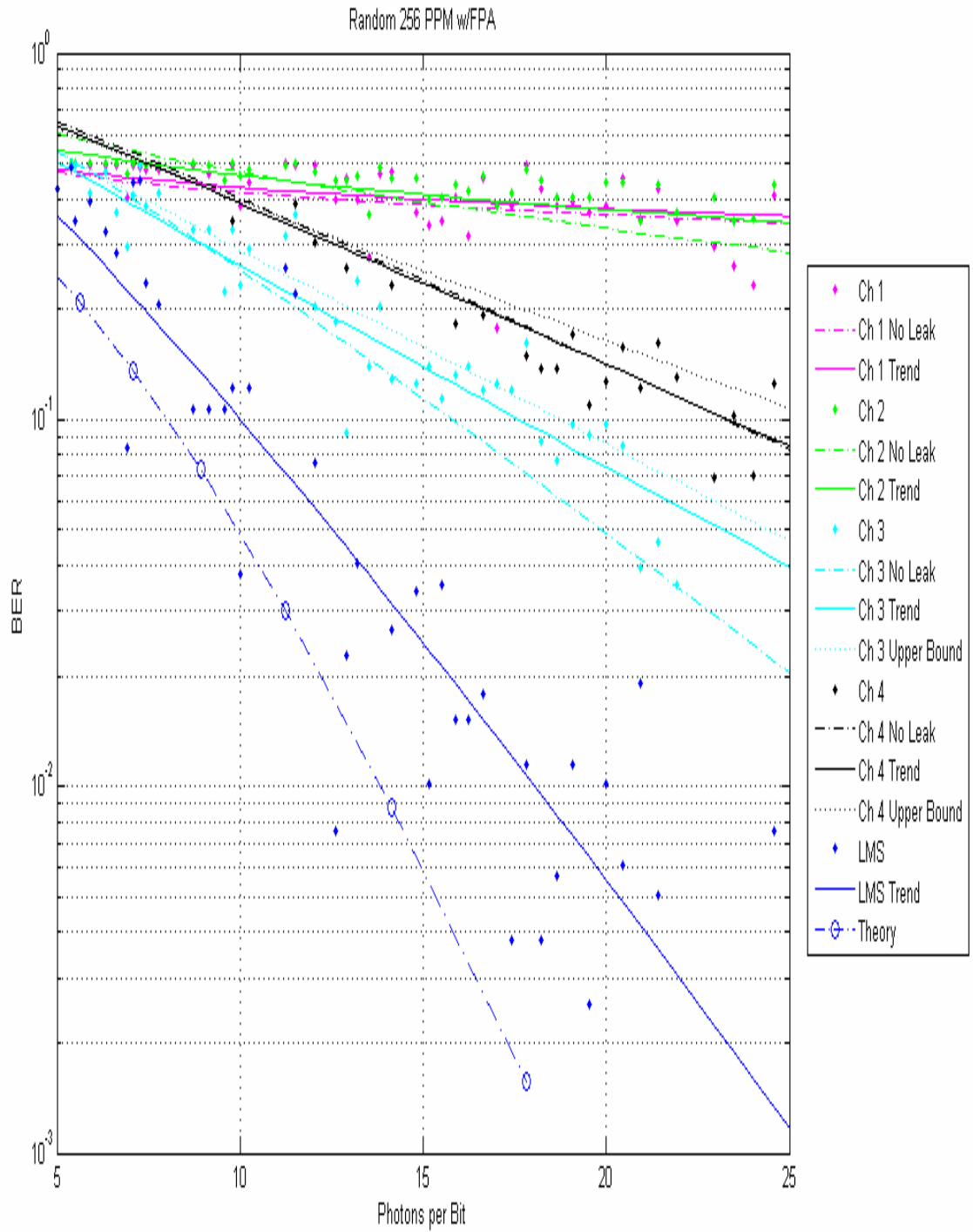


Figure 83. BER vs. photons per bit for 256 PPM with the FPA.

BER performance vs. K_s was far less impressive than what was achieved with the single detector. Nevertheless, as shown in Figure 76, BER performance as a function of SNR comes within roughly 1.0 to 1.5 dB of theory. This is a strong indication that LMS combining performance is good. If the same LMS algorithm and the same signal to noise ratios had existing in a hypothetical FPA system with better detectors characterized by Eq. (6.2.4) (the equation for the single detector) instead of Eq. (6.2.6), BER performance vs. K_s would have been far more impressive.

Figures 78 and 79 illustrate the potential of an FPA with shot noise limited detectors and amplifiers. LMS combining, in Figure 79, achieves a level of performance within 0.5 photons per bit of the theoretical limit for an envelope detection receiver. Furthermore, it is clearly seen that the FPA's performance easily exceeds that of any single channel, including the strongest, channel 3. The ability to coherently combine signals in the focal plane allows the system to achieve BER performance close to that achievable without atmospheric turbulence using a single detector and is a significant improvement over single detector performance in the presence of turbulence. In particular, most of the signal photon energy is successfully recovered as evidenced by the closeness of the BER curve to the theoretical curve. Performance is close to what would have been achieved if no turbulence were present and if all of the signal photons had landed on a single detector instead of on the multiple detection elements of the FPA.

Similar comments hold for the case of 32 PPM illustrated in the next six figures.

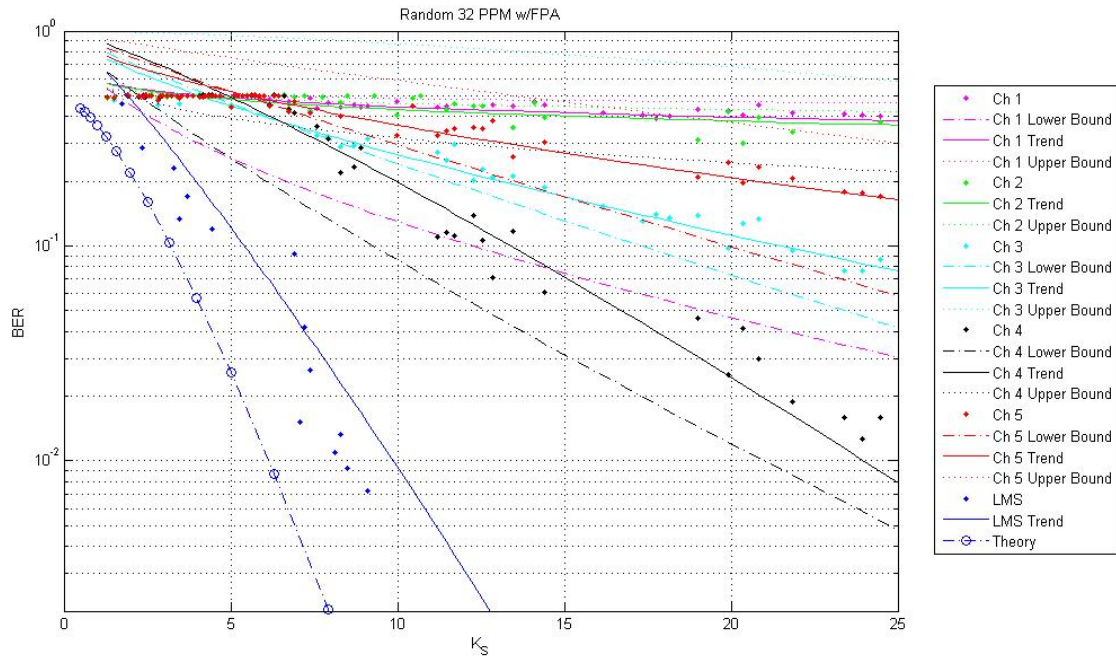


Figure 84. BER vs. K_S for a hypothetical shot-noise-limited FPA (32 PPM).

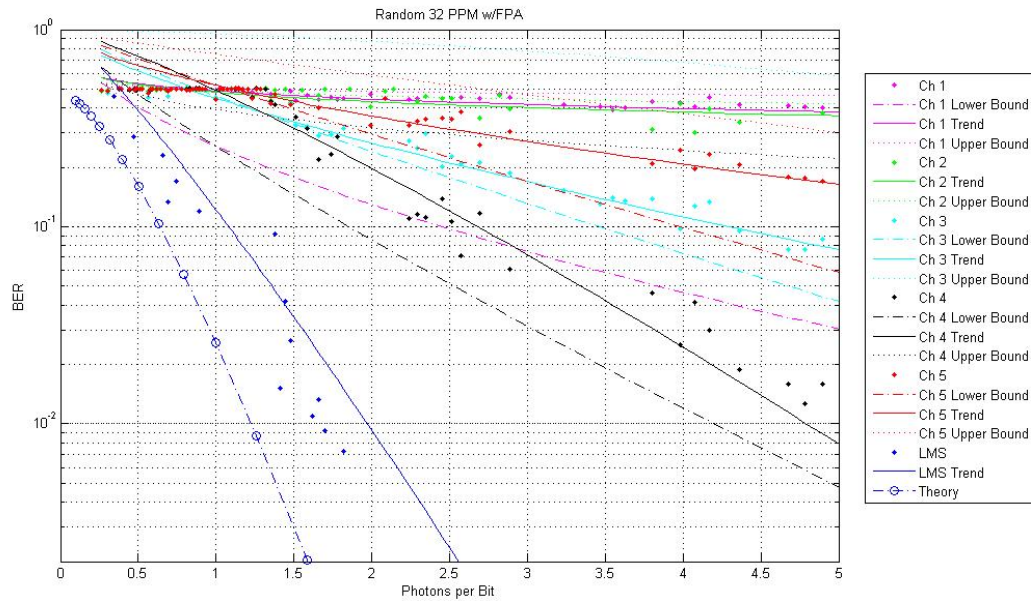


Figure 85. BER vs. photons per bit for a hypothetical shot-noise-limited FPA (32 PPM).

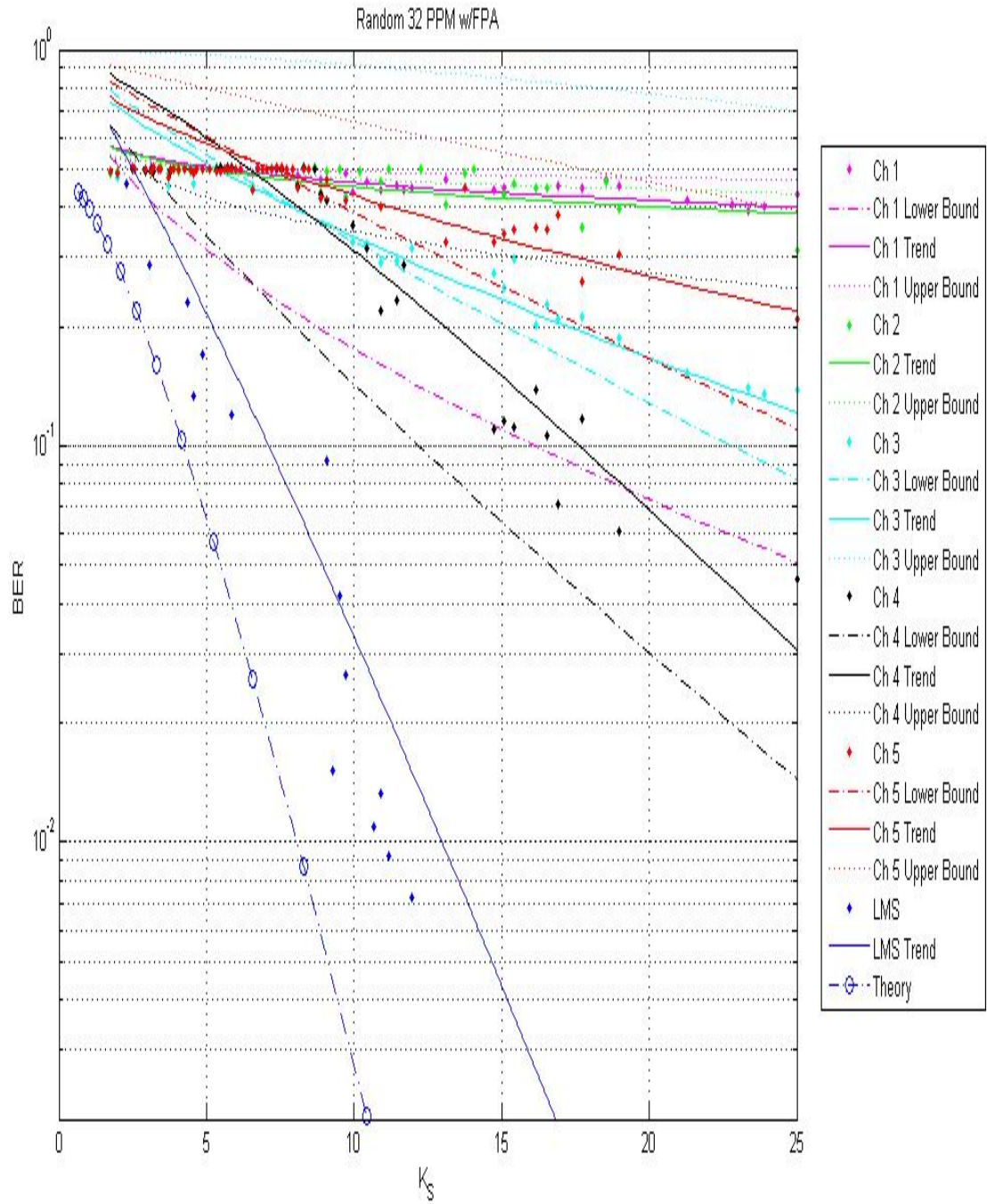


Figure 86. BER vs. K_s for a hypothetical FPA with the same noise characteristics as the single detector (32 PPM).

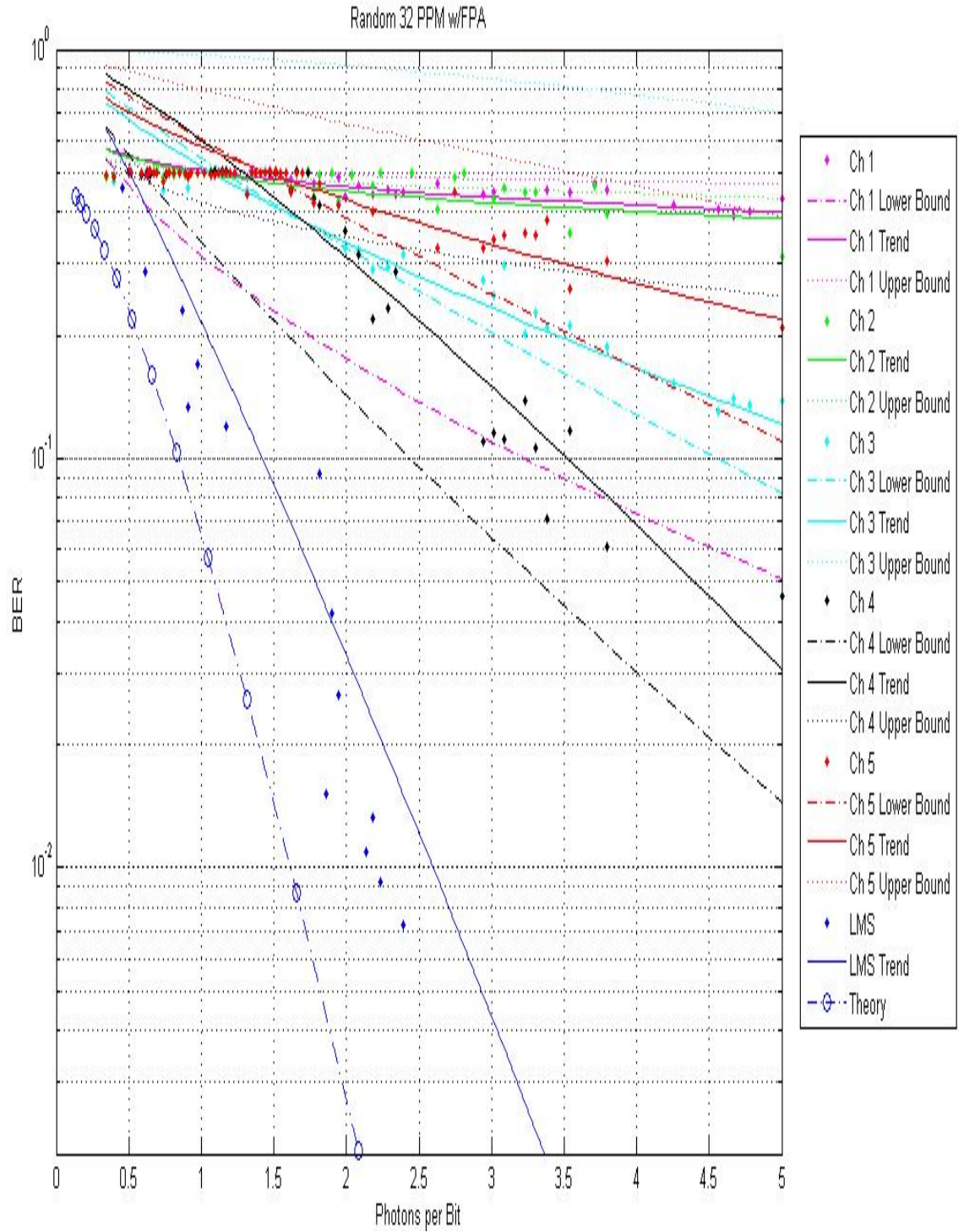


Figure 87. BER vs photons per bit for a hypothetical FPA with the same noise characteristics as the single detector (32 PPM).

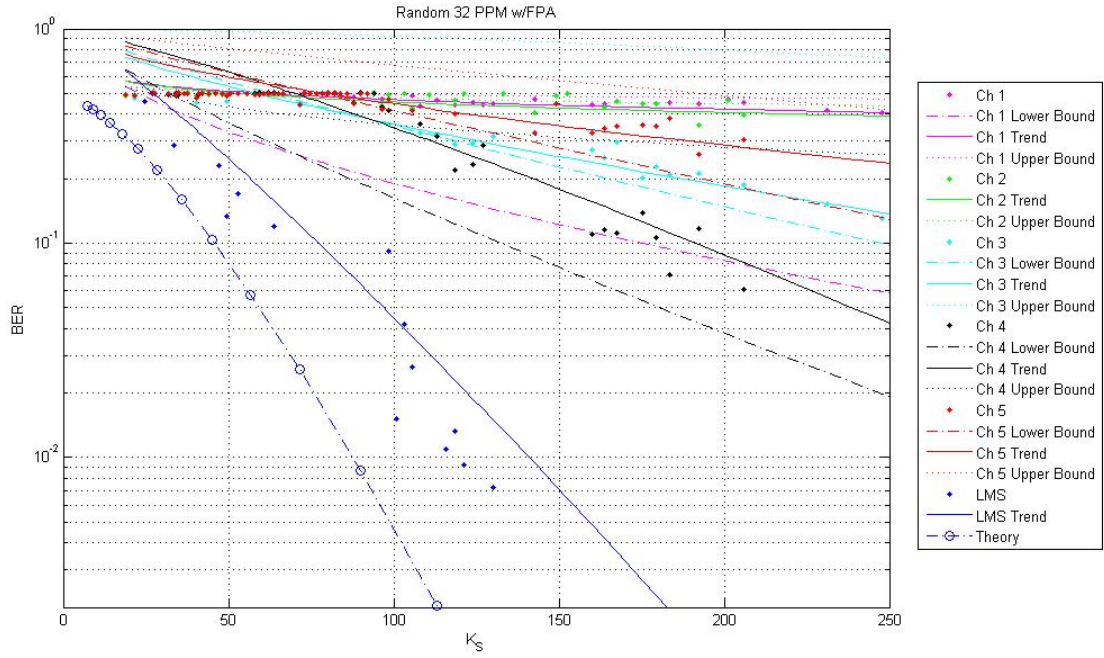


Figure 88. BER vs K_S actually achieved with the FPA (32 PPM).

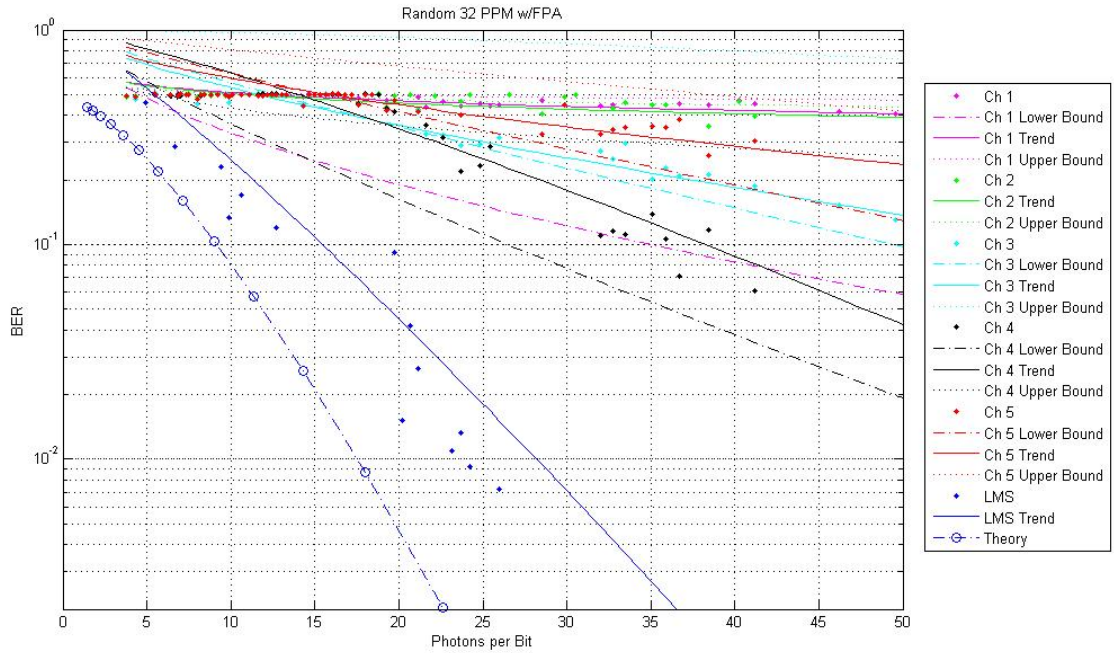


Figure 89. BER vs photons per bit actually achieved with the FPA (32 PPM).

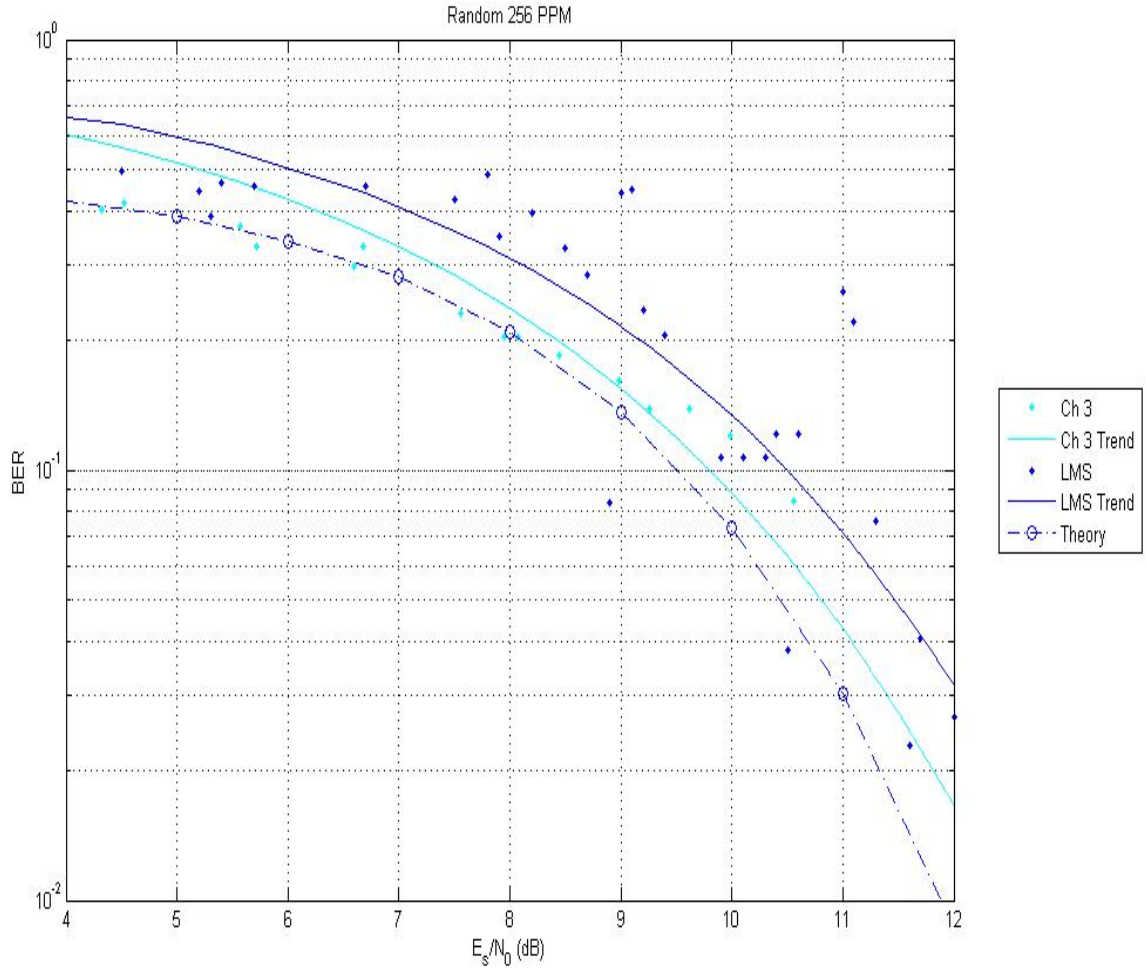


Figure 90. BER vs. E_s/N_0 for 256 PPM: Illustration of single channel performance vs. combined channel. Single channel BER is plotted vs. single channel SNR and not against combined channel SNR as in previous figures.

Figure 90 shows BER as a function of SNR for 256 PPM. Performance of a single channel (channel 3) is compared to that of the combined channel. It is important to notice that single channel BER is plotted vs. individual channel SNR and not against combined channel SNR as in previous figures. One observes a 0.2 dB loss due to the

leakage term in channel 3 illustrated by the horizontal difference between the theoretical (blue-dotted curve) and the experimental channel 3 curve (cyan). The combining loss for this case is 0.8 dB when comparing the experimental channel 3 curve with the experimental combined channel curve (continuous blue curve).

In summary, high quality low noise receiver electronics and/or a sufficiently powerful local oscillator laser operating at or near the shot noise limit, used in conjunction with a focal plane array, will enable bit error rate performance close to that achievable in the absence of atmospheric turbulence. Again, combined channel performance is within less than 0.5 photons per bit of the theoretical performance curve, indicating that much of the signal energy scattered across the five FPA elements used in combining can be recovered. These results, like those for 256 PPM, indicate that coherently combining the outputs of the FPA channels permits very good reconstruction of the signal that would have existed if no turbulence had been present and if all photons had been hitting this detector.

6.3 Conclusions

The FPA experimental results indicate the ability of the focal plane array in conjunction with the LMS algorithm to coherently and constructively add signals together in order to achieve performance very close to that of a single detector without any atmospheric turbulence. The BER curves for the combined output of the FPA channels are close to

the theoretical curves for a single channel system at the same SNR as the FPA combined channel. The spreading of signal energy in the detector focal plane caused by atmospheric turbulence can therefore be corrected by means of coherent combining of the focal plane array channels using the LMS algorithm with decision feedback. Thus, the FPA-based receiver in the laboratory achieves the goal of compensating for the phase distortions introduced at the aperture plane of the system and continues to achieve robust performance in the face of such distortions. The large improvement in BER performance of the complete FPA over even the best single channel in the detector plane illustrates the fact that it is necessary to use spatial combining to recover from signal spreading in the focal plane induced by phase errors in the aperture plane caused by atmospheric turbulence.

The strong performance of the single detector system provides real world laboratory verification of the ability to detect weak, low SNR signals in a reliable fashion as well. Given the ability to receive such weak signals when high quality electronics are used, it is believed that the use of high quality low thermal noise amplifiers and circuits in an FPA system would yield very good results in the presence of turbulence. The plots of hypothetical FPA performance assuming noise characteristics similar to those of the single detector clearly illustrates the potential that the FPA has, although the actual FPA suffered from significant thermal noise degradation.

Chapter 7: Conclusions and Future Directions

7.1 Summary of key results

The overall goal of this thesis is the development of a new set of optically coherent receiver technologies for deep space laser communications. Specifically, there are two key challenges to deep space communications that were addressed:

1. The weak signal challenge. Although optical communications offer significantly narrower beamwidths and reduced free space losses compared to RF communications, free space losses will still necessitate good weak signal reception capability.
2. The atmospheric turbulence challenge. Atmospheric turbulence distorts the incoming plane wave from the spacecraft laser transmitter. These phase front distortions result in significant scattering of the received signal in the focal plane of the detector. The challenge is to gather the spatial components of this signal and re-assemble them to recover the original signal.

The Coherent Optical Receiver Experiment rises to these challenges. Specifically, the new technologies developed here

1. Provide the ability to receive telemetry successfully at signal levels as low as one bit per photon, demonstrated experimentally with a single detector and shown to be feasible theoretically for a focal plane array.

2. Provide coherent combining capabilities to enable recovery of a signal scattered over the focal plane field by atmospheric turbulence. In the laboratory, signals have been recovered to within almost 1 dB of the original signal using a focal plane array with LMS combining.

Strong evidence for the ability to achieve good weak signal performance with an FPA in the presence of atmospheric turbulence has been given, illustrating the FPA's true potential using higher quality detector electronics and, possibly, a stronger local oscillator to achieve shot noise limited performance. Theoretical calculations describing the performance of an optically coherent focal plane array receiver have been performed, and experimental results provide good verification of these calculations, and the LMS algorithm in its application to the problem of optically coherent deep space communications has been analyzed for convergence and performance properties.

In summary, the system designed presented herein has been shown, through both analysis and experiments, to have the potential to achieve good BER performance in the presence of weak signals and in the presence of distorted focal plane signal fields that result from phase errors due to atmospheric turbulence. The technology developed here can enable fundamental gains in the performance of deep space communication links by overcoming both weak signal limitations and signal degradation due to atmospheric phase errors.

7.2 Future directions

The results presented here are highly promising, and there are many very productive directions for future research.

Coherent combining algorithms for focal plane arrays are an area of further research. In a realistic space communications system utilizing PPM, the relative positions of the spacecraft and the ground receiver may change by many optical wavelengths between PPM pulses. Since no signal exists during the inter-pulse intervals, tracking of the spacecraft signal phase poses a challenge under some operating conditions, depending on the degree of pulse-to-pulse phase coherence present.

The use of coding would unleash the true potential of the focal plane array receiver. Uncoded bit error rates below 10^{-1} are achievable using a shot-noise-limited receiver running at slightly less than one photon per bit. Corresponding coded bit error rates would be significantly lower, enabling very robust digital communications at low signal levels. The ability to communicate reliably at such weak signal levels would greatly extend the utility and usefulness of optical communications for deep space, especially in light of the severe demands generated by deep space link budgets.

Finally, more future work would include theoretical derivation of a quantum-optimum receiver for turbulence-degraded optical fields. The objective would be to find the structure and performance of the receiver that yields the best performance in the reception of signals that are described quantum mechanically. The principles of statistical detection and estimation theory would be used with the laws of quantum

mechanics taken into account in order to show the configuration of the best quantum receiver for the specific case treated in this thesis (Hoversten 1967, Hoversten, Harger and Halme, 1970, Kennedy and Hoversten, 1968, Helstrom, 1967, Helstrom, Liu and Gorden, 1970).

Appendix: A Mathematical Characteristic of Leakage

In Chapter 5, it was argued that the spatial distribution of the parasitic leakage signal in the noise slots differed from the spatial distribution of the desired signal in the signal slots. The arguments of Chapter 5 relied extensively on two vectors. The vector **a**, which was the nominal signal vector in the signal slots, is not parallel to the vector **b**, which is the corresponding leakage vector in the noise slots. In this appendix, physical evidence for this assertion is presented.

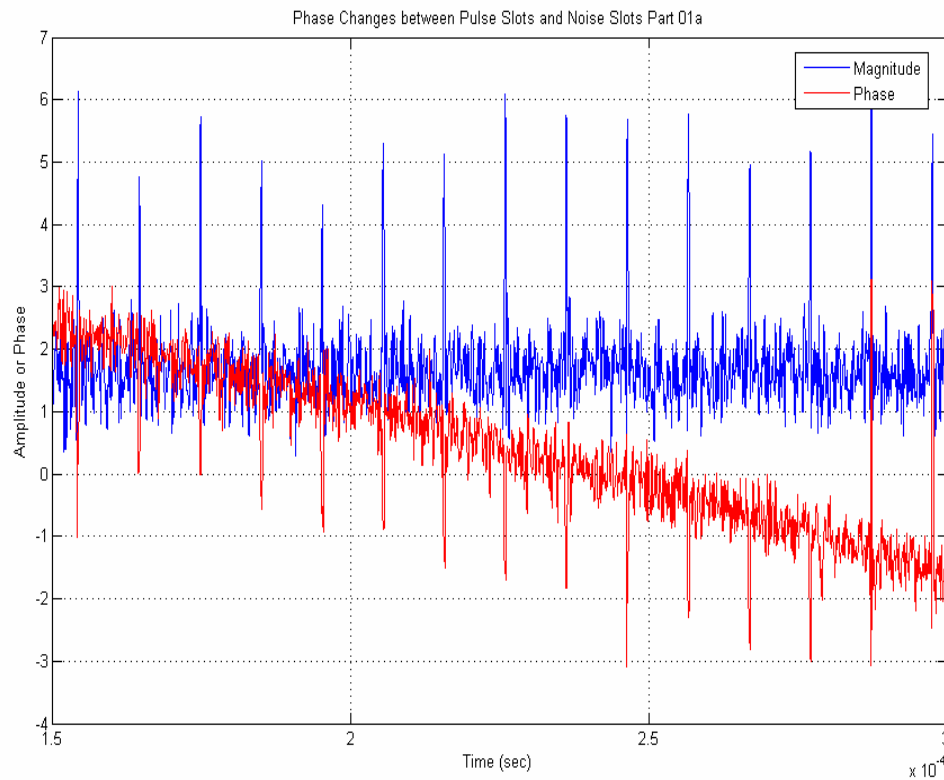


Figure A1. Magnitude and phase of the first channel plotted as a function of time.

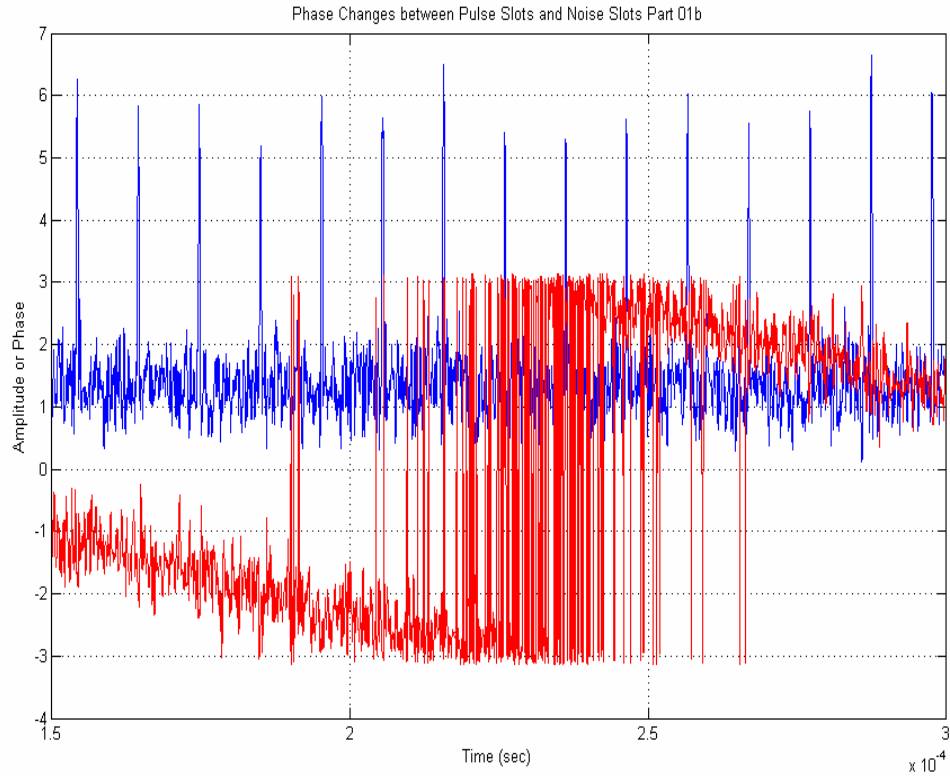


Figure A2. Magnitude and phase of the second channel plotted as a function of time. The phase behavior of this channel differs greatly from that of the first.

The above two figures illustrate phase and magnitude as a function of time for two FPA channels. Magnitudes are shown in blue, with sharp peaks denoting the presence of PPM signal pulses. Phases, in radians, are shown in red. Although only two channels are shown, these two figures already illustrate a property of the leakage.

First, in Figure A.1, there are sharp changes in phase that are clearly synchronized with the PPM pulses. This indicates that in channel 1, the phase of the channel changes depending on whether we are in a signal slot or in a noise slot. If the vectors **a** and **b** truly are parallel, then one would have to be a constant multiple of the other. This means

that the phase “spikes” in all channels would need to be of the same magnitude. For example, if a phase spike of 2 radians occurs in the first channel during a PPM pulse, the same phase spike of 2 radians would have to occur in all other channels if **a** and **b** truly are parallel. Yet it is observed that the behavior of the phase in Figure A.2 differs greatly from the behavior in Figure A.1. In particular, the first four phase spikes in Figure A.1 that occur synchronously with the first four PPM signal pulses are absent in Figure A.2. There is a phase spike between the fourth and fifth PPM signal pulses in Figure A.2, however, that is not present in Figure A.1. Given the very different phase behaviors of these two channels, one cannot conclude that the first and second elements of vectors **a** and **b** are at all parallel. Hence, the physical evidence forces one to assume that **a** and **b** are not parallel in general, and such disparate phase behavior has been observed many times in the laboratory.

Bibliography

- G. L. Abbas, V. W. S. Chan, and T. K. Yee. 1985. "A Dual Detector Optical Heterodyne Receiver for Local Oscillator Noise Suppression." *Journal of Lightwave Technology*, vol. LT-3, no. 5, pp. 1110-1122.
- S. B. Alexander. 1987. "Design of Wide-Band Optical Heterodyne Balanced Mixer Receivers." *Journal of Lightwave Technology*, vol. LT-5, no. 4, pp. 523-537.
- S. B. Alexander and B. Stephen. 1997. *Optical Communication Receiver Design*. The Society of Photo-Optical Instrumentation Engineers, SPIE Tutorial Texts in Optical Engineering vol. TT22.
- L. Andrews. 2004. *Field Guide to Atmospheric Optics*. SPIE Field Guides, vol. FG02.
- Y. Arimoto. 2004. "Study on Laser Communication Demonstration Equipment at the International Space Station." *Journal of the National Institute of Information and Communications Technology of Japan*, vol. 51, nos. 1/2.
- N. Benvenuto and G. Cherubini. 2002. *Algorithms for Communications Systems and their Applications*. West Sussex, England: John Wiley and Sons.
- S. Betti, G. de Marchis, and E. Iannone. 1995. *Coherent Optical Communication Systems*. New York: Wiley.
- V. W. S. Chan, L. L. Jeromin, and J.E. Kaufmann. 1983. "Heterodyne LASERCOM Systems Using GaAs Lasers for ISL Applications." *IEEE ICC '83*, pp. 1201-1207, Boston, Massachusetts.

- V. W. S. Chan. 1987. "Space coherent optical communications systems, an introduction." *J Lightwave Technol.*, vol 5, n. 4, pp. 633-637.
- V. W. S. Chan. 2000. "Optical space communications." *IEEE J. Sel. Top. Quantum Electron.*, vol. 6, n. 6, pp. 959-975.
- V. W. S. Chan. 2003. "Optical satellite networks." *J. Lightwave Technol.*, vol. 21, n. 11, pp. 2811-2827.
- W Chapman and M. Fitzmaurice. 1991. "Optical Space-to-Ground Link Availability Assessment and Diversity Requirements." *Free-Space Laser Communication Technologies III*, Los Angeles, California, SPIE.
- C. Chen. *Deep Space Optical Communications*. October 2005. Chapter 2- Link and System Design, ed. by H. Hemmati. Deep Space Communications and Navigation Systems, edited by the Deep Space Communications and Navigation Systems Center of Excellence, Editor-in-Chief, J. Y. Yuen. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.
- S. Chen, A. K. Samangan, B. Mulgrew, and L. Hanzo. May 2001. "Adaptive Minimum-BER linear Multiuser Detection." *Proc. (ICASSP '01), IEEE International Conference on Acoustics, Speech, and Signal Processing*. vol. 4, pp. 7-11.
- C. Arthur Clarke. 1945. *Extra-terrestrial Relays*. UK: Wireless World.
- R. T. Compton. 1988. *Adaptive Antennas*. New Jersey: Prentice Hall.
- F. Davidson, and X. Sun. 1989. "Slot Clock Recovery in Optical PPM Communication Systems with Avalanche Photodiode Detectors." *IEEE Transactions on Communications*, vol. 37, no. 11, pp. 1164-1172.

- F. Davidson and X. Sun. 1991. "Bandwidth Requirements for Direct Optical Communication Receivers with PPM Signaling." *Free-Space Laser Communication Technologies III, SPIE*, vol. 1417.
- O. E. DeLange. 1972. "Wideband optical communications: Part II-frequency division multiplexing." *Proc. IEEE*, vol. 58, n. 10, pp. 1683-1690.
- R. M. Gagliardi and Sherman Karp, *Optical Communications*. 1995. Wiley Series in Telecommunications and Signal Processing, 2nd edition, ed. by John Proakis, New York: John Wiley & Sons.
- L. C. Godara. August 1997. "Application of Antenna Arrays to Mobile Communications, Part II: Beam-Forming and Direction-of-Arrival Considerations." *Proceedings of the IEEE*, vol. 85, no. 8.
- J. W. Goodman. 2005. *Introduction to Fourier Optics*. Englewood, Colorado: Roberts & Co.
- F. E. Goodwin. 1967. "A 3.39-micron infrared optical heterodyne communications system." *IEEE J. Quantum Electron.*, QE, vol. 3, n. 11, pp. 524-531.
- S. Haykin. 1991. *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice Hall.
- C. W. Helstrom. March 1967. "Detectability of coherent optical signals in a heterodyne receiver." *J. Opt. Soc. Am.*, vol. 57, pp. 353-361.
- C. W. Helstrom, J. W. S. Liu, and J. P. Gordon. October 1970. "Quantum-mechanical communication theory, *Proc. IEEE*, vol. 58, no. 10, pp. 1578-1598.
- P. S. Henry and S. D. Personick. 1990. *Coherent Lightwave Communications*. New York: IEEE Press.

- C. A. Holt. 1978. *Chapter 20 – Feedback Amplifiers in Electronic Circuits: Digital and Analog.*, New York: John Wiley and Sons.
- P. W. Hooijmans. 1994. *Coherent Optical Receiver Design.*, New York: John Wiley and Sons.
- M. L. Honig and D. G. Messerschmitt. 1984. *Adaptive filters: structures, algorithms, and applications.* Boston: Kluwer.
- E. V. Hoversten. 1967. “The atmosphere as an optical communication channel.” *1967 IEEE Internatl. Conv. Rec.*, pt. 11, vol. 15, pp. 137-145.
- E. V. Hoversten, R. O. Harger and S. J. Halme. 1970. “Communication Theory for the Turbulent Atmosphere.” *Proc. IEEE*, vol. 58, no. 10, pp. 1626-1649.
- J. L. Hullet and T. V. Muoi. October 1976. “A Feedback Receive Amplifier for Optical Transmission Systems.” *IEEE Transactions on Communications*, vol. COM-24, no. 10, pp. 1180-1185.
- J. Katz. August 15, 1982. “2.5 Bit/Detected Photon Demonstration Program: Phase II and III Experimental Results.” *The Telecommunications and Data Acquisition Progress Report 42-70*, Jet Propulsion Laboratory, Pasadena, California, pp. 95-104.
- J. Katz. 1986. “Average Power Constraints in AlGaAs Semiconductor Lasers Under Pulse-Position-Modulation Conditions.” *Optics Communications*, vol. 56, no. 5, pp. 330-333.
- K. Kaufmann. Retrieved in May 2005. “Choosing your detector, Signal-to-noise ratio, active area, and cost requirements drive the choice of detectors.” <http://oemagazine.com/fromTheMagazine/mar05/tutorial.html>.

- R. S. Kennedy and E. V. Hoversten. September 1968. "On the atmosphere as an optical communication channel." *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 716-725.
- K. Kikuchi and K. Katoh. 2002a. "Differential detection of single modulation sideband for ultra-dense optical frequency-division multiplexed systems systems." *Electron. Lett.*, vol. 38, n. 17, pp. 980-981.
- K. Kikuchi and K. Katoh. 2002b. "Optical heterodyne receiver for selecting densely frequency division multiplexed signals." *Electron. Lett.*, vol. 38, n.16, pp. 283-285.
- T. Kuri and K. Kitayama. 2002. "Optical heterodyne detection of millimeter-wave-band radio-on-fiber signals with a remote dual-mode local light source." *IEEE Trans. Microwave Theory & Techn.*, vol. 49, n. 10, pp. 2025-2029.
- T. Kuri and K. Kitayama. 2003. "Optical heterodyne detection technique for densely multiplexed millimeter-wave-band radio-on-fiber systems." *J. Lightwave Technol.*, vol. 21. n. 12, pp. 3167-3179.
- J. R. Lesh, J. Katz, H. H. Tan, and D. Zwillinger. December 15, 1981. "2.5-Bit/Detected Photon Demonstration Program: Description, Analysis and Phase I Results." *The Telecommunications and Data Acquisition Progress Report 42-66*, Jet Propulsion Laboratory, Pasadena, California, pp. 115-132.
- J. R. Lesh. 1982. "Power Efficient Communications for Space Applications, in International Telemetry Conference." *International Foundation for Telemetry*, vol. XVIII.
- J. R. Lesh and D.L. Robinson. 1986. "A Cost-Performance Model for Ground-Based Optical Communications Receiving Telescopes." *TDA PR 42-87*, pp. 56-64.

- J. R. Lesh. October 2005. *Deep Space Optical Communications*. Chapter 1-Introduction. Ed. by H. Hemmati. Deep Space Communications and Navigation Systems, edited by the Deep Space Communications and Navigation Systems Center of Excellence, Editor-in-Chief, J. Y. Yuen. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.
- Lightwave Electronics. 1988. *Users Manual* Series 120 Diode Pump Ring Laser, Lightwave Electronics Corporation, Mountain View, CA.
- J. H. McElroy, N. McAvoy, E. H. Johnson, J. J. Degnan, F. E. Goodwin, D. M. Henderson, T. A. Nussmeier, L. S. Stokes, B. J. Peyton, and T. Flattau. February 1977. "CO₂ Laser Communications Systems for Near Earth Space Applications." *Proc. IEEE*, vol. 65, no. 2, pp. 221-251.
- A. MacGregor and B. Dion. 1991. "39 Photons/bit Direct Detection Receiver at 810 nm, BER=1E-6, 60 Mb/s, QPPM." *Free-Space Laser Communications Technologies III*, *SPIE*, vol. 1417.
- M. Muñoz Fernández, and V. A. Vilnrotter. 2004. "Performance analysis and preliminary experimental verification of a coherent optical receiver for PPM signals in the presence of atmospheric turbulence." *Proc. 5th International Conference on Space Optics (ICSO 2004)*, Toulouse, France. Ed. by B. Warmbein. ESA SP-554, Noordwijk, Netherlands: ESA Publications Division, ISBN 92-9092-865-4, 2004, p. 411-418.

- M. Muñoz Fernández, and V. A. Vilnrotter. 2005a. "Optical system for reception of coherently detected PPM signals in the presence of atmospheric turbulence." Accepted to *Journal Acta Astronautica, International Academy of Astronautics* (IAA).
- M. Muñoz Fernández and Victor A Vilnrotter. 2005b. "Optical System for Reception of Coherently Detected PPM Signals in the Presence of Atmospheric Turbulence," *Interplanetary Network Progress Report, IPN* vol. 42-161. Joseph H. Yuen Editor in Chief. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.
- M. Muñoz Fernández, and V. A. Vilnrotter. 2005c. "Coherent Optical Receiver for PPM Signals under Atmospheric Turbulence." *Proc. IEEE Aerospace Conference*, pp. 1-8.
- T. A. Nussmeier, F. E. Goodwin, and J. E. Zavin. 1974. "A 10.6- μm terrestrial communication link." *IEEE J. Quantum Electron.*, QE vol. 10, n. 2, pp. 230-235.
- T. Okoshi, and K. Kikuchi. 1988. *Coherent Optical Fiber Communications*. Tokyo: KTK Scientific.
- B. M. Oliver. 1961. "Signal-to-Noise Ratios in Photoelectric Mixing." *Proc. IEEE*, vol. 49, n. 12, pp. 1960-61.
- G. R. Osche. 2002. *Optical Detection Theory for Laser Applications*. Hoboken, New Jersey: Wiley Series in Pure and Applied Optics, ed. by Bahaa E. A. Saleh.
- A. Papoulis. 1991. *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill.

- B. J. Peyton. 1972. A. J. Dinardo, G. M. Kanischak, F. R. Arams, R. A. Lange, and E. W. Sard. "High-sensitivity receiver for infrared laser communications." *IEEE J. Quantum Electron.*, QE vol. 8, n. 2, pp. 252-263.
- J. Pierce. Dec 1978. "Optical Channels: Practical Limits with Photon Counting." *IEEE Transactions on Communications*, vol. COM-26, n. 12, pp. 1819–1821.
- J. G. Proakis. 2000. *Digital Communications*, 4th edition. New York: McGraw Hill.
- E. R. Rochat, Dandliker, K. Haroud, R. H. Czichy, U. Roth, D. Constantini, and R. Holzner. 2001. "Fiber amplifiers for coherent space communication." *IEEE J. Sel. Top. Quantum Electron.*, vol. 7, n. 1, pp. 64-80.
- M. Ross. 1966. *Laser Receivers*. New York: Wiley.
- S. Ryu. 1995. *Coherent Lightwave Communication Systems*. Boston: Artec House.
- A. Sayed. 2003. *Fundamentals of Adaptive Filtering*. New Jersey: John Wiley and Sons.
- M. K. Simon, S. Hinedi, and W. Lindsey. 1995. *Digital Communication Techniques, Signal Design and Detection*. Upper Saddle River, New Jersey: PTR Prentice Hall.
- M. K. Simon. 2002. *Probability Distributions Involving Gaussian Random Variables, A Handbook for Engineers and Scientists*. Norwell, Massachusetts: Kluwer Academic Publishers.
- W.S. Streifer. 1989. *Characteristics of High Power GaAlAs Laser Diodes Useful for Space Applications* in Optical Space Communication, SPIE.
- X. Sun, F. M. Davidson, and C. T. Field. July 1990. "50 Mbps Free-Space Laser Direct Detection Laser Diode Optical Communication System with Q = 4 PPM Signaling."

- Free-Space Laser Communication Technologies II, SPIE*, vol. 1218, pp. 385-395, ed. By David L. Begley and Bernard D. Seery.
- X. Sun and F. Davidson. 1990. "Word Timing Recovery in Direct Detection Optical PPM Communication Systems with Avalanche Photodiodes Using a Phase Lock Loop." *IEEE Transactions on Communications*, vol. 38, no. 5, pp. 666-673.
- E. A. Swanson, J. C. Livas, and R. S. Bondurant. Feb. 1994. "High Sensitivity Optically Preamplified Direct Detection DPSK Receiver with Active Delay-Line Stabilization." *IEEE Photonics Technology Letters*, vol. 6, n. 2, pp. 263–265.
- Y. Takasaki, M. Tanaka, N. Maeda, K. Yamashita, and K. Nagano. April 1976. "Optical Pulse Formats for Fiber Optic Digital Communications." *IEEE Transactions on Communications*, vol. COM-24, no. 4, pp. 404-413.
- J. Taylor. M. Muñoz Fernández, A. Bolea Alamanac, and Kar-Ming Cheung. October 2001. "*Deep Space 1 Telecommunications.*" Design and Performance Summary Series, Deep Space Communications and Navigation Systems Center of Excellence (DESCANSO), Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, U.S.A. http://descanso.jpl.nasa.gov/DPSummary/All_Article2.pdf
- J. Taylor, A. Makovsky, A. Barbieri, R. Tung, P. Estabrook and A. Gail Thomas. 2005. "Mars Exploration Rover Telecommunications." Design and Performance Summary Series, Deep Space Communications and Navigation Systems Center of Excellence (DESCANSO), Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, U.S.A.
- W. Van Driel. 2004. "Frequency protection for 21st century instruments." *Exploring the Cosmic Frontier-Astrophysical Instruments for the 21st Century*, Berlin.

- V. A. Vilnrotter, E. R. Rodemich, and H. H. Tan. 1986. "A Synchronization Technique for Optical PPM Signals." TDA PR 42-87, pp. 24-31.
- V. A. Vilnrotter, Rodemich, E.R.; Dolinar, S.J., Jr. March 1992. "Real-time combining of residual carrier array signals using ML weight estimates." *IEEE Transactions on Communications*, vol. 40, n. 3, pp. 604–615.
- V. A. Vilnrotter and Srinivasan. July 2002. "Adaptive detector arrays for optical communications receivers." *IEEE Transactions on Communications*, vol. 50, n. 7, pp. 1091-1097.
- A. J. Viterbi. 1966. *Principles of Coherent Communication*. New York: McGraw-Hill.
- B. Widrow and S. D. Stearns. 1985. *Adaptive Signal Processing*. Englewood Cliffs, New Jersey: Prentice Hall Signal Processing Series, ed. by A. Oppenheim.
- A. D. Wyner. Nov 1988. "Capacity and Error Exponent for the Direct Detection Photon Channel-Part I." *IEEE Transactions on Information Theory*, vol. 34, n. 6, pp. 1449-1961.
- Y. Yamamoto and H. A. Haus. October 1986. "Preparation, measurement and information capacity of optical quantum states." *Reviews of Modern Physics*, vol. 58, n. 4.
- Yamamoto, Y. 1980. "Receiver Performance Evaluation of Various Digital Optical Modulation-Demodulation Systems in the 0.5-10 micron Wavelength Region." *IEEE Journal of Quantum Electronics*, vol. 16, n. 11, pp. 1251-1259.
- A. Yariv. 1997. *Optical Electronics in Modern Communications*, 5th edition. New York: Oxford University Press.